

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 6, 2004, 00:06:13 ; Search time 145.506 Seconds  
(without alignments)  
11537.137 Million cell updates/sec

Title: US-10-626-173-1  
Perfect score: 3025  
Sequence: 1 gtcaccacgcgtccggcc.....aaaaaaaaaaggcgccgc 3025

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA.\*  
1: /cgn2\_6/ptodata/2/ina/5A COMB.seq.\*  
2: /cgn2\_6/ptodata/2/ina/5B COMB.seq.\*  
3: /cgn2\_6/ptodata/2/ina/6A COMB.seq.\*  
4: /cgn2\_6/ptodata/2/ina/6B COMB.seq.\*  
5: /cgn2\_6/ptodata/2/ina/PTCUS COMB.seq.\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3025	100.0	3025	US-09-291-839-1	Sequence 1, Appli
2	3025	100.0	3025	US-09-458-457-1	Sequence 1, Appli
3	3025	100.0	3025	US-09-947-199A-1	Sequence 1, Appli
4	2505	82.8	2505	US-09-291-839-3	Sequence 3, Appli
5	2505	82.8	2505	US-09-458-457-3	Sequence 3, Appli
6	2505	82.8	2505	US-09-947-199A-3	Sequence 3, Appli
7	2010	66.4	3026	US-09-458-457-7	Sequence 7, Appli
8	2010	66.4	3026	US-09-947-199A-7	Sequence 7, Appli
9	1892.2	62.6	2505	US-09-458-457-9	Sequence 9, Appli
10	1892.2	62.6	2505	US-09-947-199A-9	Sequence 9, Appli
11	427.6	14.1	740	US-09-833-381-1151	Sequence 1151, Ap
12	427.4	14.1	616	US-09-833-381-1149	Sequence 1149, Ap
13	414.6	13.7	736	US-09-833-381-1150	Sequence 1150, Ap
14	184.2	6.1	304	US-09-833-381-1147	Sequence 1147, Ap
15	84.4	2.8	186	US-09-833-381-1148	Sequence 1148, Ap
16	84	2.8	14770	US-09-220-132-30	Sequence 30, Appl
17	78.8	2.6	2894	US-09-620-312D-95	Sequence 95, Appl
18	78.8	2.6	3606	US-10-164-595-37	Sequence 37, Appl
19	78.6	2.6	3454	US-09-082-059-1	Sequence 1, Appli
20	75.4	2.5	4297	US-09-972-115A-3	Sequence 3, Appli
21	75.4	2.5	4493	US-09-972-115A-1	Sequence 1, Appli
22	74.2	2.5	384	US-09-972-115A-9	Sequence 9, Appli
23	74.2	2.5	1069	US-09-833-381-841	Sequence 841, App
24	74.2	2.5	3394	US-09-427-154-1	Sequence 1, Appli
25	74.2	2.5	3498	US-09-350-982C-4	Sequence 4, Appli
26	74.2	2.5	3797	US-09-696-668-1	Sequence 1, Appli
27	74.2	2.5	3816	US-09-696-668-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1

US-09-291-839-1  
; Sequence 1, Application US/09291839A  
; Patent No. 6261818  
; GENERAL INFORMATION:  
; APPLICANT: Raju, Jeyaseelan  
; TITLE OF INVENTION: NOVEL CARP PROTEIN AND NUCLEIC ACID MOLECULES AND USES THEREFOR  
; FILE REFERENCE: MNI-068  
; CURRENT APPLICATION NUMBER: US/09/291.839A  
; CURRENT FILING DATE: 1999-04-14  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 3025  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (48)..(2552)  
US-09-291-839-1

Query Match		100.0%;	Score 3025;	DB 3;	Length 3025;
Best Local Similarity		100.0%;	Pred. No. 0;		
Matches 3025;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	GTGACCCACCGCTCCGGCCCTGGAGAAAGAGAAACTTATTAATGGGAAATATA	60		
Db	1	GTGACCCACCGCTCCGGCCCTGGAGAAAGAGAAACTTATTAATGGGAAATATA	60		
QY	61	AATCTAGACCAACCCAACTTGTACTGATGAATGGAAGAAAGTCAAGTGAATCATATG	120		
Db	61	AATCTAGACCAACCCAACTTGTACTGATGAATGGAAGAAAGTCAAGTGAATCATATG	120		
QY	121	TTATCACAATAGAAGATTAGAAGTCACTGCAGATCAAGGAAAGAACTCAAGAAC	180		
Db	121	TTATCACAATAGAAGATTAGAAGTCACTGCAGATCAAGGAAAGAACTCAAGAAC	180		
QY	181	TAAGGATATATTTGGCTCTGATGAAGCCCTCAGTAAAGTCAATTTAAATCCGCACTG	240		
Db	181	TAAGGATATATTTGGCTCTGATGAAGCCCTCAGTAAAGTCAATTTAAATCCGCACTG	240		
QY	241	AAATGGGCTCTCTACTTCTCAATTTATGTTGCAATTTGTGGAGGCAAGAAATCATATTC	300		
Db	241	AAATGGGCTCTCTACTTCTCAATTTATGTTGCAATTTGTGGAGGCAAGAAATCATATTC	300		
QY	301	GAATCTTTATGTAAGGGCTCGCCCATCTCGACTGACAGAAATGTTTACAGCCT	360		
Db	301	GAATCTTTATGTAAGGGCTCGCCCATCTCGACTGACAGAAATGTTTACAGCCT	360		
QY	361	TGCATTTAGCATTTTACAGGATTAATGCAAGATTGATCACTTCTCTCTTCACAGTGGAG	420		

Sequence 5, Appli  
Sequence 3, Appli  
Sequence 138, App  
Sequence 140, App  
Sequence 138, App  
Sequence 140, App  
Sequence 35, Appl  
Sequence 36, Appl  
Sequence 35, Appl  
Sequence 36, Appl  
Sequence 35, Appl  
Sequence 36, Appl  
Sequence 32, Appl  
Sequence 32, Appl  
Sequence 34, Appl  
Sequence 34, Appl

Db	361	TGCAATTTAGCAGTTTACAAGGATAATGCAGAAATTCATCTCTGCTTCCACAGTGGAG	420	Db	1441	TCCATGAGATTAATGGCTCAGGTTCTTTTGGGAAAGATATATAAAGGACGATGCAGAAATA	1500
Qy	421	CTGATATACAGCAGGTTCGATAGGTTGGCTCAGTGCCTCCATATTCCTCAATAGCTG	480	Qy	1501	AAATAGTGGCTATAAAACGTTTATCGAGCCAAATACCTACTGCTCCAACTCAGATGTGGATA	1560
Db	421	CTGATATACAGCAGGTTCGATAGGTTGGCTCAGTGCCTCCATATTCCTCAATAGCTG	480	Db	1501	AAATAGTGGCTATAAAACGTTTATCGAGCCAAATACCTACTGCTCCAACTCAGATGTGGATA	1560
Qy	481	GCCACCTAGAGGCTGCTGATGTGCTGTTGCAACATGAGCTAAATGCAATATTCAGATG	540	Qy	1561	TGTTTTGGCGAGAGGTGTCATTTCTCTGCCAGCTCAATCATCCCTGGTAAATTCAGTTTG	1620
Db	481	GCCACCTAGAGGCTGCTGATGTGCTGTTGCAACATGAGCTAAATGCAATATTCAGATG	540	Db	1561	TGTTTTGGCGAGAGGTGTCATTTCTCTGCCAGCTCAATCATCCCTGGTAAATTCAGTTTG	1620
Qy	541	CAGTTTTTTTCACTCCATTTGCATATTTGAGCGCTACTATGGACATGAACAGTAACTGGCC	600	Qy	1621	TGGGTGCTTGTGTAATCATCCAGCCAGTTTGCCATTGTCACTCAATACATATCAGGGG	1680
Db	541	CAGTTTTTTTCACTCCATTTGCATATTTGAGCGCTACTATGGACATGAACAGTAACTGGCC	600	Db	1621	TGGGTGCTTGTGTAATCATCCAGCCAGTTTGCCATTGTCACTCAATACATATCAGGGG	1680
Qy	601	TTCTTTTGAATTTGGTGTGATGTAATGTAAGTGGTGAAGTTGGAGATAGACCCCTCC	660	Qy	1681	GTTCTCTGTTCTCCCTCCTTCATGAGCAGAAGAGGATTTCTTGATTTGCACTCTAAATTA	1740
Db	601	TTCTTTTGAATTTGGTGTGATGTAATGTAAGTGGTGAAGTTGGAGATAGACCCCTCC	660	Db	1681	GTTCTCTGTTCTCCCTCCTTCATGAGCAGAAGAGGATTTCTTGATTTGCACTCTAAATTA	1740
Qy	661	ACCTAGCATCTGCAAAAGGATTTGGAATTTGCAAAATCTCTTGATGGAAGGACGCA	720	Qy	1741	TTATTTGCGATAGATGTTGCCAAAGGATGAGTACCTTCAACCTGCACACGCAATTA	1800
Db	661	ACCTAGCATCTGCAAAAGGATTTGGAATTTGCAAAATCTCTTGATGGAAGGACGCA	720	Db	1741	TTATTTGCGATAGATGTTGCCAAAGGATGAGTACCTTCAACCTGCACACGCAATTA	1800
Qy	721	AAGCAGATGTGAATGCTCAAGATAATGAAGACCATGTCCCACTCCATTTCTGTTCTGAT	780	Qy	1801	TACATCGTGAATTTGAAACAGTCACAATATTTCTTCTATGAGGATGGGCATGCTGTGGTG	1860
Db	721	AAGCAGATGTGAATGCTCAAGATAATGAAGACCATGTCCCACTCCATTTCTGTTCTGAT	780	Db	1801	TACATCGTGAATTTGAAACAGTCACAATATTTCTTCTATGAGGATGGGCATGCTGTGGTG	1860
Qy	781	TTGGACACCATGATATAGTTAAAGTATCTGCTGCAAAAGTGAATTTGGAAGTTCAACCTCATG	840	Qy	1861	CAGATTTTGGAGAAATCAAGATTTCTACAGTCTCTGGATGAAGACACATGACAAAACAAC	1920
Db	781	TTGGACACCATGATATAGTTAAAGTATCTGCTGCAAAAGTGAATTTGGAAGTTCAACCTCATG	840	Db	1861	CAGATTTTGGAGAAATCAAGATTTCTACAGTCTCTGGATGAAGACACATGACAAAACAAC	1920
Qy	841	TTGTTTAATATCTATGAGATACCCCTTACACTGGCATGCTACAAATGGCAAAATTTGAAG	900	Qy	1921	CTGGAAACCTCCGTTGGATGGCTCTGAGGTGTTTCAAGCAGTGCATCGGTACACCATCA	1980
Db	841	TTGTTTAATATCTATGAGATACCCCTTACACTGGCATGCTACAAATGGCAAAATTTGAAG	900	Db	1921	CTGGAAACCTCCGTTGGATGGCTCTGAGGTGTTTCAAGCAGTGCATCGGTACACCATCA	1980
Qy	901	TTGCCAAGGAAATCATCCAAATATCAGGAACAGAAAGTCTGACTAAGGAAACATCTTCA	960	Qy	1981	AAGCAGATGTCTTACAGTATGCTGTGTCTGTGGGAAATTTCTCACTGGGAAATTTCCAT	2040
Db	901	TTGCCAAGGAAATCATCCAAATATCAGGAACAGAAAGTCTGACTAAGGAAACATCTTCA	960	Db	1981	AAGCAGATGTCTTACAGTATGCTGTGTCTGTGGGAAATTTCTCACTGGGAAATTTCCAT	2040
Qy	961	GTGAAACAGCTTTTCAATAGTGTGTTGACCTATGGCAAGAGCATTTGACCTAGTCAAAATTC	1020	Qy	2041	TCGCTCATCTCAAGCCAGCGCTGGCGAGCAGACATGGCTTACCACCAATCAGACCTC	2100
Db	961	GTGAAACAGCTTTTCAATAGTGTGTTGACCTATGGCAAGAGCATTTGACCTAGTCAAAATTC	1020	Db	2041	TCGCTCATCTCAAGCCAGCGCTGGCGAGCAGACATGGCTTACCACCAATCAGACCTC	2100
Qy	1021	TTCTTGATCAGATGTCATAAATCAATCAACCAAGGAGGATGGGCACACTGGATTAC	1080	Qy	2101	CAATTGGCTATTTCCATTTCCCAAGCCCATATCATCTCTGCTGATACGAGGTTGAAACGCT	2160
Db	1021	TTCTTGATCAGATGTCATAAATCAATCAACCAAGGAGGATGGGCACACTGGATTAC	1080	Db	2101	CAATTGGCTATTTCCATTTCCCAAGCCCATATCATCTCTGCTGATACGAGGTTGAAACGCT	2160
Qy	1081	ACTCTGCTTGCTACACGCTCAGTTTCCCTGCTGTTTCTTACTTGATTAATGAGCTG	1140	Qy	2161	GTCCTGAGGAGAGCCGAAATTTTCTGAAGTTGTCATGAAGTTAGAGAGTGTCTCTGCA	2220
Db	1081	ACTCTGCTTGCTACACGCTCAGTTTCCCTGCTGTTTCTTACTTGATTAATGAGCTG	1140	Db	2161	GTCCTGAGGAGAGCCGAAATTTTCTGAAGTTGTCATGAAGTTAGAGAGTGTCTCTGCA	2220
Qy	1141	ATATGAATCTAGTGGCTTGATGCCAGCAGCTCTAGTGGTGAABAAAGATGAGCAGACAT	1200	Qy	2221	ACATTGAGCTGATGTCTCTCATCAAGTAAACAGCAGTGGGTCTCTCTCACCTTCTTCTT	2280
Db	1141	ATATGAATCTAGTGGCTTGATGCCAGCAGCTCTAGTGGTGAABAAAGATGAGCAGACAT	1200	Db	2221	ACATTGAGCTGATGTCTCTCATCAAGTAAACAGCAGTGGGTCTCTCTCACCTTCTTCTT	2280
Qy	1201	GTTTGATGTGGGCTTATGAAAAGGAGGATGATGATGTCACACTCTCTGAAAGCATTATA	1260	Qy	2281	CTTCTGATTTGCTGTGAAACCGGGAGGACCTGGCCGAGTCAATGGCAGCATTAAGAA	2340
Db	1201	GTTTGATGTGGGCTTATGAAAAGGAGGATGATGATGTCACACTCTCTGAAAGCATTATA	1260	Db	2281	CTTCTGATTTGCTGTGAAACCGGGAGGACCTGGCCGAGTCAATGGCAGCATTAAGAA	2340
Qy	1261	AGAGACCAAGATGAATTTGCCCTGTAATGAATTTCTCAGCCTGGAGAGATGGCTCTCT	1320	Qy	2341	GTCGTTTCCGAATTTGGAATATGCTCTAAATGCAAGGTCTCTATGCTGCTTTTCCCAAAGTG	2400
Db	1261	AGAGACCAAGATGAATTTGCCCTGTAATGAATTTCTCAGCCTGGAGAGATGGCTCTCT	1320	Db	2341	GTCGTTTCCGAATTTGGAATATGCTCTAAATGCAAGGTCTCTATGCTGCTTTTCCCAAAGTG	2400
Qy	1321	ATGTGCTGTTTCCATCACCTTGGGGAAGATTAAGACATGACAAAAGAGAGGACAGATA	1380	Qy	2401	CTGGACAATATTTCTCTCAAGGTCTCTTGGAGGAGATGAAGAGAGTCTTCAATACA	2460
Db	1321	ATGTGCTGTTTCCATCACCTTGGGGAAGATTAAGACATGACAAAAGAGAGGACAGATA	1380	Db	2401	CTGGACAATATTTCTCTCAAGGTCTCTTGGAGGAGATGAAGAGAGTCTTCAATACA	2460
Qy	1381	TTCTCTCTTAAGAGCTGGAATTTGCCCTTCAATTTTCCATTTTCCATTTTCCATTTTCCAT	1440	Qy	2461	CACCCTTGCACAAATATGCTTATGCCGATCCCATGAGCTCAATGCAATTTTCAATCTT	2520
Db	1381	TTCTCTCTTAAGAGCTGGAATTTGCCCTTCAATTTTCCATTTTCCATTTTCCATTTTCCAT	1440	Db	2461	CACCCTTGCACAAATATGCTTATGCCGATCCCATGAGCTCAATGCAATTTTCAATCTT	2520
Qy	1441	TCCATGAGATTAATTTGGCTCAGGTTCTTTTGGGAAAGTATATAAAGGACGATGCAGAAATA	1500	Qy	2521	GCCGAAATAGTAGCAGCTTTGAGGACAGCAGCTGACAGCATTCGCGGTATACCTTAAGGAG	2580
Db	1441	TCCATGAGATTAATTTGGCTCAGGTTCTTTTGGGAAAGTATATAAAGGACGATGCAGAAATA	1500	Db	2521	GCCGAAATAGTAGCAGCTTTGAGGACAGCAGCTGACAGCATTCGCGGTATACCTTAAGGAG	2580

```
Qy 2581 AGTTTTTCCCGAAGCTGACAGCAAGATTCCACACCGCAAGCTGGCTTCCAACTATA 2640
Db 2581 AGTTTTTCCCGAAGCTGACAGCAAGATTCCACACCGCAAGCTGGCTTCCAACTATA 2640
Qy 2641 ACATTTTACTCTCAAAGGCTCTCTTAAATGGGCTTGTTTTTACTTGTCCTATTAAATTC 2700
Db 2641 ACATTTTACTCTCAAAGGCTCTCTTAAATGGGCTTGTTTTTACTTGTCCTATTAAATTC 2700
Qy 2701 CCACATATTAGCGGCTTTGGATTGTGCTTAAGCAATTAATGCAAAAGAACCAAGACA 2760
Db 2701 CCACATATTAGCGGCTTTGGATTGTGCTTAAGCAATTAATGCAAAAGAACCAAGACA 2760
Qy 2761 GAATGTATATGAAGAAATGTTTTTAATTTTGTAATTAATAAATAAATTTAGATCGTTACT 2820
Db 2761 GAATGTATATGAAGAAATGTTTTTAATTTTGTAATTAATAAATAAATTTAGATCGTTACT 2820
Qy 2821 TGGAAATGGAGCCTAAAGTCTGTGGTGACAGATAATAATATATGTTTCTGGGCTGAAT 2880
Db 2821 TGGAAATGGAGCCTAAAGTCTGTGGTGACAGATAATAATATATGTTTCTGGGCTGAAT 2880
Qy 2881 ATGTAGACTTTGTGTTGACAGCTATGGTTTATTTCTTGAACATTTGTTCAATTTCTTTT 2940
Db 2881 ATGTAGACTTTGTGTTGACAGCTATGGTTTATTTCTTGAACATTTGTTCAATTTCTTTT 2940
Qy 2941 CTCATTATGTTACTTCTAGTGTTCACCTCTGTGATTAAAGATTCTTTGGTGAATAGAAA 3000
Db 2941 CTCATTATGTTACTTCTAGTGTTCACCTCTGTGATTAAAGATTCTTTGGTGAATAGAAA 3000
Qy 3001 AAAAAAAAAAAAAAGGGCGGCGC 3025
Db 3001 AAAAAAAAAAAAAAGGGCGGCGC 3025
```

## RESULT 2

```
US-09-458-457-1
; Sequence 1, Application US/09458457
; Patent No. 6500654
; GENERAL INFORMATION:
; APPLICANT: Raju, Jeyaseelan
; TITLE OF INVENTION: NOVEL CARP PROTEIN AND NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: mni-068cp
; CURRENT APPLICATION NUMBER: US/09/458,457
; CURRENT FILING DATE: 1999-12-10
; EARLIER APPLICATION NUMBER: 60/111,938
; EARLIER FILING DATE: 1998-12-11
; EARLIER APPLICATION NUMBER: 09/291,839
; EARLIER FILING DATE: 1999-04-14
; NUMBER OF SEQ. ID NOS: 9
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 1
; LENGTH: 3025
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (48)..(2552)
US-09-458-457-1
```

```
Query Match 100.0%; Score 3025; DB 4; Length 3025;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3025; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTCACCCACGCGTCCGGCCCTGGAGAAAGAACTTATTAATAATGGGAATATA 60
Db 1 GTCACCCACGCGTCCGGCCCTGGAGAAAGAACTTATTAATAATGGGAATATA 60
Qy 61 AATCTAGACCAACCCAACTTGTACTGATGAGAGAAAAGTCAGTGAATCATATG 120
Db 61 AATCTAGACCAACCCAACTTGTACTGATGAGAGAAAAGTCAGTGAATCATATG 120
Qy 121 TTATCAATAGAAAGATTAGAAGATGACCTGACATCAAGAAAAAGAACTGACAGAAC 180
```

```
Db 121 TTATCAATAGAAAGATTAGAAGATGACCTGACATCAAGAAAAAGAACTGACAGAAC 180
Qy 181 TAAGGAATATATTTGGCTCTGATGAAGCCCTTCAGTAAAGTCAATTTAAATTTACCGCACTG 240
Db 181 TAAGGAATATATTTGGCTCTGATGAAGCCCTTCAGTAAAGTCAATTTAAATTTACCGCACTG 240
Qy 241 AAAATGGGCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 300
Db 241 AAAATGGGCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 300
Qy 301 GAATCTTTATGTTGAAGGGCTCCGCGCCATCTCGACTGACAAAGAAATGGATTACAGCCT 360
Db 301 GAATCTTTATGTTGAAGGGCTCCGCGCCATCTCGACTGACAAAGAAATGGATTACAGCCT 360
Qy 361 TGCATTTAGCAGTTTACAAGGATAATGCAAGATTGATCACTTCTCTCTCTCTCTCTCTCTCTCT 420
Db 361 TGCATTTAGCAGTTTACAAGGATAATGCAAGATTGATCACTTCTCTCTCTCTCTCTCTCTCTCT 420
Qy 421 CTGATATACAGCAGGTTGGATACGGTGGCCTCACTGCCCTCCATATTCCTCAATTTAGCTG 480
Db 421 CTGATATACAGCAGGTTGGATACGGTGGCCTCACTGCCCTCCATATTCCTCAATTTAGCTG 480
Qy 481 GCCACCTAGAGGCTGCTGATGCTGTTGCAACATGGAGCTAATGTCATATTTCAAGATG 540
Db 481 GCCACCTAGAGGCTGCTGATGCTGTTGCAACATGGAGCTAATGTCATATTTCAAGATG 540
Qy 541 CAGTTTTTTTTCACTCCATATTCGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
Db 541 CAGTTTTTTTTCACTCCATATTCGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
Qy 601 TTCTTTTGAATTTGGTGTGATGTAATGTAAGTGGTGAAGTTGGAGATAGACCCCTCC 660
Db 601 TTCTTTTGAATTTGGTGTGATGTAATGTAAGTGGTGAAGTTGGAGATAGACCCCTCC 660
Qy 661 ACCTAGCATCTGCAAAAGGATTCTTGAATATTCGAAACTCTTGATGGAAGAGGACGACA 720
Db 661 ACCTAGCATCTGCAAAAGGATTCTTGAATATTCGAAACTCTTGATGGAAGAGGACGACA 720
Qy 721 AAGCAGATGTAATGCTAAGATAATGAAGACATGTCCTCCATCCATTTCTTCTCGAT 780
Db 721 AAGCAGATGTAATGCTAAGATAATGAAGACATGTCCTCCATCCATTTCTTCTCGAT 780
Qy 781 TTGGACACCATGATATAGTTAAGTATCTGCTCAAGTGAATTTGGAAGTTCAACCTCATG 840
Db 781 TTGGACACCATGATATAGTTAAGTATCTGCTCAAGTGAATTTGGAAGTTCAACCTCATG 840
Qy 841 TTGTTAATATCTATGGAGATACCCCTTACACCTGGCATGCTACAATGGCAAAATTTGAAG 900
Db 841 TTGTTAATATCTATGGAGATACCCCTTACACCTGGCATGCTACAATGGCAAAATTTGAAG 900
Qy 901 TTGCCAAGGAAATCATCCAAATATCAGGAACAGAAAGTCTGACTAAGGAACATCTTCA 960
Db 901 TTGCCAAGGAAATCATCCAAATATCAGGAACAGAAAGTCTGACTAAGGAACATCTTCA 960
Qy 961 GTGAAACAGCTTTTTCATAGTGTGTTACCTATGGCAAGAGCATTGACCTAGTCAAAATTC 1020
Db 961 GTGAAACAGCTTTTTCATAGTGTGTTACCTATGGCAAGAGCATTGACCTAGTCAAAATTC 1020
Qy 1021 TTCTTGATCAGAATGTCAAAACATCAACACCAAGAGGATGGGCACTGGATTAC 1080
Db 1021 TTCTTGATCAGAATGTCAAAACATCAACACCAAGAGGATGGGCACTGGATTAC 1080
Qy 1081 ACTCTGCTTCTACCGGTCACATTCGGCTGTTTCTGATTTCTTACTGGATAATGGAGCTG 1140
Db 1081 ACTCTGCTTCTACCGGTCACATTCGGCTGTTTCTGATTTCTTACTGGATAATGGAGCTG 1140
Qy 1141 ATATGAATCTAGTGGCTTTGATCCAGCAGGCTTAGTGGTGAAGAAAGATGAGCAGACAT 1200
Db 1141 ATATGAATCTAGTGGCTTTGATCCAGCAGGCTTAGTGGTGAAGAAAGATGAGCAGACAT 1200
Qy 1201 GTTTGATGCTGGGCTTATGAAAAAGGCGATGATCCCATCTGCATCTCTGAGGCACTTATA 1260
Db 1201 GTTTGATGCTGGGCTTATGAAAAAGGCGATGATCCCATCTGCATCTCTGAGGCACTTATA 1260
```

Db 1201 GTTTGATGTGGCTTATGAAAAAGGCGATGATGCCATTGTTCACACTCTCTGAAGCAATTATA 1260  
QY 1261 AGAGACCACAGATGAATTCCTGCTTAATGAATATTTCTCAGCCTGGAGGAGATGGCTCCT 1320  
Db 1261 AGAGACCACAGATGAATTCCTGCTTAATGAATATTTCTCAGCCTGGAGGAGATGGCTCCT 1320  
QY 1321 ATGTGCTCTCCATCACCCTTGGGAGAGATTAAGACATGACAAAAAGAGAGAGCAGATA 1380  
Db 1321 ATGTGCTCTCCATCACCCTTGGGAGAGATTAAGACATGACAAAAAGAGAGAGCAGATA 1380  
QY 1381 TTCTCTCTTAAGAGCTGAATTCCTTCAATTCCTTCAATTCCTTCAATTCCTTCAATTCCTT 1440  
Db 1381 TTCTCTCTTAAGAGCTGAATTCCTTCAATTCCTTCAATTCCTTCAATTCCTTCAATTCCTT 1440  
QY 1441 TCCATGAGATTAATGGCTCAGGTTCCTTTGGGAGAGATTAAGAGGAGATGACAGAAATA 1500  
Db 1441 TCCATGAGATTAATGGCTCAGGTTCCTTTGGGAGAGATTAAGAGGAGATGACAGAAATA 1500  
QY 1501 AAATAGTGGCTATAAAACGTTATCGAGCCAAATACCTACTGCTCCAAGTCAAGATGGGATA 1560  
Db 1501 AAATAGTGGCTATAAAACGTTATCGAGCCAAATACCTACTGCTCCAAGTCAAGATGGGATA 1560  
QY 1561 TGTGTTGCCGAGAGTGCATTCCTGCGAGCTCAATCATCCCTGGTAAATTCAGTTTG 1620  
Db 1561 TGTGTTGCCGAGAGTGCATTCCTGCGAGCTCAATCATCCCTGGTAAATTCAGTTTG 1620  
QY 1621 TGGGTGCTTGTGAATGATCCAGCCAGTTTGCCTTGTGCTCAATCATCATATATCAGGGG 1680  
Db 1621 TGGGTGCTTGTGAATGATCCAGCCAGTTTGCCTTGTGCTCAATCATCATATATCAGGGG 1680  
QY 1681 GTTCTCTGTTCTCCCTCTTATGAGCAGAGAGATTCCTTGAATTCGAGTCTTAATTA 1740  
Db 1681 GTTCTCTGTTCTCCCTCTTATGAGCAGAGAGATTCCTTGAATTCGAGTCTTAATTA 1740  
QY 1741 TTATTCAGTACAGATTCCTCAAGGATGGAGTACCTTCAACCTGACACAGCCAATTA 1800  
Db 1741 TTATTCAGTACAGATTCCTCAAGGATGGAGTACCTTCAACCTGACACAGCCAATTA 1800  
QY 1801 TACATCGTACCTGAACAGTCAATATTCCTTCTATGAGGATGGGATGGGATGGG 1860  
Db 1801 TACATCGTACCTGAACAGTCAATATTCCTTCTATGAGGATGGGATGGGATGGG 1860  
QY 1861 CAGATTTTGGAGATCAAGATTCCTACAGTCTCTGGATGAAGACAAATGACAAACAC 1920  
Db 1861 CAGATTTTGGAGATCAAGATTCCTACAGTCTCTGGATGAAGACAAATGACAAACAC 1920  
QY 1921 CTGGGAACCTCCGTTGGATGGCTCCTGAGGTGTTACGCACTCGGTACACCATCA 1980  
Db 1921 CTGGGAACCTCCGTTGGATGGCTCCTGAGGTGTTACGCACTCGGTACACCATCA 1980  
QY 1981 AAGCAGATGCTTACGATGCTGCTGTCTGTGGGAAATTCCTACCTGGGAAATTCAT 2040  
Db 1981 AAGCAGATGCTTACGATGCTGCTGTCTGTGGGAAATTCCTACCTGGGAAATTCAT 2040  
QY 2041 TGGCTCATCTCAAGCAGCGGCTCGGAGCAGACATGGCTTACCACCAATCAGACCTC 2100  
Db 2041 TGGCTCATCTCAAGCAGCGGCTCGGAGCAGACATGGCTTACCACCAATCAGACCTC 2100  
QY 2101 CATTGGCTTATTCATTCACAGCCCATATCATCTCTGATACGAGGTTGAGACGCAT 2160  
Db 2101 CATTGGCTTATTCATTCACAGCCCATATCATCTCTGATACGAGGTTGAGACGCAT 2160  
QY 2161 GTCTTGAAGAGAGACCCGAAATTCCTGAAGTTGTCATGAAGTTAGAGAGTCTCTGCA 2220  
Db 2161 GTCTTGAAGAGAGACCCGAAATTCCTGAAGTTGTCATGAAGTTAGAGAGTCTCTGCA 2220  
QY 2221 ACATTGAGCTGATGCTCTCTGATCAAGTAAACAGAGTGGTCTCTCTCACTTCTTCTT 2280  
Db 2221 ACATTGAGCTGATGCTCTCTGATCAAGTAAACAGAGTGGTCTCTCTCACTTCTTCTT 2280  
QY 2281 CTTCTGATTCCTGAGTACCGGGAGGACCTGGCCGAGTCAATGCGCAGCATTAAGAA 2340  
Db 2281 CTTCTGATTCCTGAGTACCGGGAGGACCTGGCCGAGTCAATGCGCAGCATTAAGAA 2340

QY 2341 GTCTGTTTCCGAATTCGAATATGCTCTAAATGCAAGGTCTCTATGCTTGTGTTGCCAAGTG 2400  
Db 2341 GTCTGTTTCCGAATTCGAATATGCTCTAAATGCAAGGTCTCTATGCTTGTGTTGCCAAGTG 2400  
QY 2401 CTGGACAATAATTCCTCTCAAGGTCTGCTTTGGAGGAGATGAAAAAGAGTCTTCAATACA 2460  
Db 2401 CTGGACAATAATTCCTCTCAAGGTCTGCTTTGGAGGAGATGAAAAAGAGTCTTCAATACA 2460  
QY 2461 CACCCATATGCAAAATATGGCTATGATCCGATCCCATGAGCTCAATGCAATTTTCAATCTT 2520  
Db 2461 CACCCATATGCAAAATATGGCTATGATCCGATCCCATGAGCTCAATGCAATTTTCAATCTT 2520  
QY 2521 GCCGAATATGATAGCAGCTTTGAGGACAGCAGCTGACAGCATTCGGGCTATACCTAAGGAG 2580  
Db 2521 GCCGAATATGATAGCAGCTTTGAGGACAGCAGCTGACAGCATTCGGGCTATACCTAAGGAG 2580  
QY 2581 AGTTTTCCTCCGAACTGACAGCAACGATTCGAACCGCAAGCTGGCTTCCAACTATA 2640  
Db 2581 AGTTTTCCTCCGAACTGACAGCAACGATTCGAACCGCAAGCTGGCTTCCAACTATA 2640  
QY 2641 ACATTTTACTCTCAAGGTCTCTTAAATTTGGGCTTGTCTTAACTTGTCTTAAATTC 2700  
Db 2641 ACATTTTACTCTCAAGGTCTCTTAAATTTGGGCTTGTCTTAACTTGTCTTAAATTC 2700  
QY 2701 CCCACTATTAGCAGCTTTGGATTTGCTTAAGGAATAATATGCAAAAGAACCAAGACA 2760  
Db 2701 CCCACTATTAGCAGCTTTGGATTTGCTTAAGGAATAATATGCAAAAGAACCAAGACA 2760  
QY 2761 GAATGATATGAAGATTTGTTTAAATTTGTAATTTAAATTTAAATTTAGATCGTACT 2820  
Db 2761 GAATGATATGAAGATTTGTTTAAATTTGTAATTTAAATTTAAATTTAGATCGTACT 2820  
QY 2821 TGGAAATGGAGCCCTAAGTCTGCTGAGCAGATAAATATATGTTTCTGGGCTGAAT 2880  
Db 2821 TGGAAATGGAGCCCTAAGTCTGCTGAGCAGATAAATATATGTTTCTGGGCTGAAT 2880  
QY 2881 ATGTAGATTTGTTTACAGCAGCTATGGGTTTATTTCTTAGAACATTTGTTCTTTT 2940  
Db 2881 ATGTAGATTTGTTTACAGCAGCTATGGGTTTATTTCTTAGAACATTTGTTCTTTT 2940  
QY 2941 CTCATTATGTTACTTCTAGTGTTCACCTCTGCTGATTAAGATTTCTTTGGTGAATAGAAA 3000  
Db 2941 CTCATTATGTTACTTCTAGTGTTCACCTCTGCTGATTAAGATTTCTTTGGTGAATAGAAA 3000  
QY 3001 AAAAAAAGGAGGCGCGCGC 3025  
Db 3001 AAAAAAAGGAGGCGCGCGC 3025

## RESULT 3

US-09-947-199A-1

; Sequence 1, Application US/09947199A

; Patent No. 6660490

; GENERAL INFORMATION:

; APPLICANT: Raju, Jeyaseelan

; TITLE OF INVENTION: NOVEL CARB PROTEIN AND NUCLEIC ACID MOLECULES AND USES

; TITLE OF INVENTION: THEREFOR

; FILE REFERENCE: MNI-068CP2

; CURRENT APPLICATION NUMBER: US/09/947,199A

; PRIOR FILING DATE: 2001-09-05

; PRIOR APPLICATION NUMBER: US 60/111,938

; PRIOR FILING DATE: 1998-12-11

; PRIOR APPLICATION NUMBER: US 09/291,839

; PRIOR FILING DATE: 1999-04-14

; PRIOR APPLICATION NUMBER: US 09/458,457

; PRIOR FILING DATE: 1999-12-10

; NUMBER OF SEQ ID NOS: 9

; SOFTWARE: Patent In Ver. 2.0

; SEQ ID NO 1

; LENGTH: 3025

; TYPE: DNA

; ORGANISM: Homo sapiens

;  
; NAME/KEY: CDS  
; LOCATION: (48)..(2552)  
US-09-947-199A-1

Query Match 100.0%; Score 3025; DB 4; Length 3025;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 3025; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTCGACCCAGCGTCCGGCCCTGGAGAAAGAGAAACTTATAATTAATGGGAATTATA 60  
Db 1 GTCGACCCAGCGTCCGGCCCTGGAGAAAGAGAAACTTATAATTAATGGGAATTATA 60  
Qy 61 AATCTAGACCAACCCAAACTTGTACTGATGAATGGAAAGAAAAGTCAGTGAATCATATG 120  
Db 61 AATCTAGACCAACCCAAACTTGTACTGATGAATGGAAAGAAAAGTCAGTGAATCATATG 120  
Qy 121 TTATCACAATAGAAAGANTTGAAGATGACCTGCAGATCAAGGAAAAGAACTGACAGAAC 180  
Db 121 TTATCACAATAGAAAGANTTGAAGATGACCTGCAGATCAAGGAAAAGAACTGACAGAAC 180  
Qy 181 TAAGGAATATATTGGCTCTGATGAAGCCCTTCAATTAAGTCAATTTAAATTAACCGCACTG 240  
Db 181 TAAGGAATATATTGGCTCTGATGAAGCCCTTCAATTAAGTCAATTTAAATTAACCGCACTG 240  
Qy 241 AAAATGGGCTGTCTACTTCAATTTATGTTGCAATTTGTGAGGCAAGAAATCACATATTC 300  
Db 241 AAAATGGGCTGTCTACTTCAATTTATGTTGCAATTTGTGAGGCAAGAAATCACATATTC 300  
Qy 301 GAATCTTATGTTGAAGGGCTCGGCCATCTCGACTGACAGAAATGATTTACAGCT 360  
Db 301 GAATCTTATGTTGAAGGGCTCGGCCATCTCGACTGACAGAAATGATTTACAGCT 360  
Qy 361 TGCAATTTAGCAGTTTACAAGGAAATGACAGAAATGATTTACAGTTGAGGAG 420  
Db 361 TGCAATTTAGCAGTTTACAAGGAAATGATTTACAGTTGAGGAG 420  
Qy 421 CTGATATACAGCGTTGGATGATGCTGTTGCAATGAGCTAATGCAATTTCAAGATG 540  
Db 421 CTGATATACAGCGTTGGATGATGCTGTTGCAATGAGCTAATGCAATTTCAAGATG 540  
Qy 541 CAGTTTTTTTCACTCCATTCGATATTTGCAAGCTACTATGGACATGAACAGGTAATCGCC 600  
Db 541 CAGTTTTTTTCACTCCATTCGATATTTGCAAGCTACTATGGACATGAACAGGTAATCGCC 600  
Qy 601 TTCTTTTGAATTTGGTCTGATGTAATGTAAGTGGTGAAGTTGAGATAGACCCCTCC 660  
Db 601 TTCTTTTGAATTTGGTCTGATGTAATGTAAGTGGTGAAGTTGAGATAGACCCCTCC 660  
Qy 661 ACCTAGCATCTGCAAAAAGGATTTCTTGAATATTCGAAATTCGAAATTCGAAAGGAGCA 720  
Db 661 ACCTAGCATCTGCAAAAAGGATTTCTTGAATATTCGAAATTCGAAAGGAGCA 720  
Qy 721 AAGCAGATGTAATGCTCAAGATTAATGAAGACCATGTCCTCCATTTCTGTTCTCGAT 780  
Db 721 AAGCAGATGTAATGCTCAAGATTAATGAAGACCATGTCCTCCATTTCTGTTCTCGAT 780  
Qy 781 TTGGACACCATGATATAGTTTAAGTATCTGCTGCAAAAGTGAATTTGGAAGTTCAACCTCATG 840  
Db 781 TTGGACACCATGATATAGTTTAAGTATCTGCTGCAAAAGTGAATTTGGAAGTTCAACCTCATG 840  
Qy 841 TTGTTAATATCTATGGAGATACCCCTTACCTGGCATGCTACAATGCAAAATTTGAAG 900  
Db 841 TTGTTAATATCTATGGAGATACCCCTTACCTGGCATGCTACAATGCAAAATTTGAAG 900  
Qy 901 TTGCGAGGAAATCATCCAAATATCAGGACAGAAAGTCTGACTRAGGAAACATCTTCA 960  
Db 901 TTGCGAGGAAATCATCCAAATATCAGGACAGAAAGTCTGACTRAGGAAACATCTTCA 960

Qy 961 GTGAAACAGCGTTTTTCATAGTGTCTGTACCTATGCGCAAGAGCATTGACCTAGTCAAAATTC 1020  
Db 961 GTGAAACAGCGTTTTTCATAGTGTCTGTACCTATGCGCAAGAGCATTGACCTAGTCAAAATTC 1020  
Qy 1021 TTCTTGATCAGAAATGTCTATAACATCAACCAAGAAAGGATGGGCACACTGGATTAC 1080  
Db 1021 TTCTTGATCAGAAATGTCTATAACATCAACCAAGAAAGGATGGGCACACTGGATTAC 1080  
Qy 1081 ACTCTGCTGTCTACCAAGCTGACATTTGGCTCGCTGCTGCTTCTTACTGGATATATGAGCTG 1140  
Db 1081 ACTCTGCTGTCTACCAAGCTGACATTTGGCTCGCTGCTGCTTCTTACTGGATATATGAGCTG 1140  
Qy 1141 ATATGAATCTAGTGGCTTGTGATCCAGCAGGTCTAGTGGTCAAAAAGATGAGCAGACAT 1200  
Db 1141 ATATGAATCTAGTGGCTTGTGATCCAGCAGGTCTAGTGGTCAAAAAGATGAGCAGACAT 1200  
Qy 1201 GTTTGATGTTGGCTTATGAAGAAGGCATGATCCATTTGCTACACTCTCTGAAGCATATA 1260  
Db 1201 GTTTGATGTTGGCTTATGAAGAAGGCATGATCCATTTGCTACACTCTCTGAAGCATATA 1260  
Qy 1261 AGAGACCAACAAGATGAATTTGCCCTGTAAATGAATATTTCTCAGCCTGGAGGAGATGGCTCT 1320  
Db 1261 AGAGACCAACAAGATGAATTTGCCCTGTAAATGAATATTTCTCAGCCTGGAGGAGATGGCTCT 1320  
Qy 1321 ATGTGCTGTTCATCACTTGGGGAGATTAAGAAGCATGAACAAAGAGAGAGGAGATA 1380  
Db 1321 ATGTGCTGTTCATCACTTGGGGAGATTAAGAAGCATGAACAAAGAGAGAGGAGATA 1380  
Qy 1381 TTCTCTCTCTTAAAGCTGGATTTGCCCTTCACTTTCCATCTTCAGCTCTCAGAAATTTGAGT 1440  
Db 1381 TTCTCTCTCTTAAAGCTGGATTTGCCCTTCACTTTCCATCTTCAGCTCTCAGAAATTTGAGT 1440  
Qy 1441 TCCATGAGATTAATTTGCTCAGGTTCTTTTGGGAAAGTATATATAAGGACGATCAGAAATA 1500  
Db 1441 TCCATGAGATTAATTTGCTCAGGTTCTTTTGGGAAAGTATATAAGGACGATCAGAAATA 1500  
Qy 1501 AAATAGTGGCTATATAAGCTTATCGAGCCAAATACCTACTGCTCCAGTCAGATGGGATA 1560  
Db 1501 AAATAGTGGCTATATAAGCTTATCGAGCCAAATACCTACTGCTCCAGTCAGATGGGATA 1560  
Qy 1561 TGTTTTGCCGAGAGGTGTCCATTTCTGCGCAGCTCAATCATCCCTGCGTAATTCAGTTTG 1620  
Db 1561 TGTTTTGCCGAGAGGTGTCCATTTCTGCGCAGCTCAATCATCCCTGCGTAATTCAGTTTG 1620  
Qy 1621 TGGGTGCTTCTTGAATGATCCAGCAGGTGGCTTGGCAATTTGCTCACTCAATACATATCAGGG 1680  
Db 1621 TGGGTGCTTCTTGAATGATCCAGCAGGTGGCTTGGCAATTTGCTCACTCAATACATATCAGGG 1680  
Qy 1681 GTTCTCTGTTCTCCCTCTTCATGACAGAGAGGATTTCTTGATTTGCACTTAAATTA 1740  
Db 1681 GTTCTCTGTTCTCCCTCTTCATGACAGAGAGGATTTCTTGATTTGCACTTAAATTA 1740  
Qy 1741 TTATTCAGTATGATGTCGAAAGGATGAGTACCTTCAACCTGACACAGCCAAATTA 1800  
Db 1741 TTATTCAGTATGATGTCGAAAGGATGAGTACCTTCAACCTGACACAGCCAAATTA 1800  
Qy 1801 TACATCGTACTTGAACAGTCAATATTTCTCTATGAGGATGGGATGCTGTTGGTGG 1860  
Db 1801 TACATCGTACTTGAACAGTCAATATTTCTCTATGAGGATGGGATGCTGTTGGTGG 1860  
Qy 1861 CAGATTTTGAGAAATCAAGATTTCTACGTCTCTGATGAAGACAAACATGACAAAACAAC 1920  
Db 1861 CAGATTTTGAGAAATCAAGATTTCTACGTCTCTGATGAAGACAAACATGACAAAACAAC 1920  
Qy 1921 CTGGGAACCTCCGTTGGATGGCTCTGAGGTGTTACGCGAGTGGCATCGGTACACATCA 1980  
Db 1921 CTGGGAACCTCCGTTGGATGGCTCTGAGGTGTTACGCGAGTGGCATCGGTACACATCA 1980  
Qy 1981 AAGCAGATGCTTTTCACTATGCTGCTGTGTTGGGAAATTTCTCACTGGCGAAATTTCCAT 2040  
Db 1981 AAGCAGATGCTTTTCACTATGCTGCTGTGTTGGGAAATTTCTCACTGGCGAAATTTCCAT 2040  
Qy 2041 TCGCTCATCTCAAGCCAGCGGCTGCGGCGACAGACATGCGCTTACCACCATCAGACCTC 2100

```
Db 2041 TCCTCATCTCAAGCCAGCGCTGGCGCAGCAGCATGGCTTACCACCATCAGACCTC 2100
QY 2101 CCATTGGCTATTCCATTCCCAAGCCCATATCATCTCTGCTGATACGAGGGTGGAAAGCAT 2160
Db 2101 CCATTGGCTATTCCATTCCCAAGCCCATATCATCTCTGCTGATACGAGGGTGGAAAGCAT 2160
QY 2161 GTCTCTGAAGAGACCCCGAATTTCTGAAGTTGTTCATGAAGTTAGAAGAGTGTCTCTGCA 2220
Db 2161 GTCTCTGAAGAGACCCCGAATTTCTGAAGTTGTTCATGAAGTTAGAAGAGTGTCTCTGCA 2220
QY 2221 ACATTGAGCTGATGTCTCTCTGATCAAGTAACAGAGTGGTCTCTCTACCTTCTTCTT 2280
Db 2221 ACATTGAGCTGATGTCTCTCTGATCAAGTAACAGAGTGGTCTCTCTACCTTCTTCTT 2280
QY 2281 CTCTGATTTGCTGTGAACCGGGAGGACCTGGCCGAGTCAATGGCAGCATTAAGAA 2340
Db 2281 CTCTGATTTGCTGTGAACCGGGAGGACCTGGCCGAGTCAATGGCAGCATTAAGAA 2340
QY 2341 GTCTGTTTGAATTTGAATATGCTCTTAAATGCAAGTCTCTATGCTGCTTTTGTCCAAAGTG 2400
Db 2341 GTCTGTTTGAATTTGAATATGCTCTTAAATGCAAGTCTCTATGCTGCTTTTGTCCAAAGTG 2400
QY 2401 CTGGCAATATTCCTCTCAAGTCTCTCTTTGGAGGAGATCAAGAGTCTTCAATACA 2460
Db 2401 CTGGCAATATTCCTCTCAAGTCTCTCTTTGGAGGAGATCAAGAGTCTTCAATACA 2460
QY 2461 CACCCATTGACAAATATGGCTATGTATCCGATCCCATGAGCTCAATGCAATTTTCATCTT 2520
Db 2461 CACCCATTGACAAATATGGCTATGTATCCGATCCCATGAGCTCAATGCAATTTTCATCTT 2520
QY 2521 GCGGAAATGATGAGCTTTGAGGACAGCAGCTGACAGCATTCGGCGTATACCTTAGGAG 2580
Db 2521 GCGGAAATGATGAGCTTTGAGGACAGCAGCTGACAGCATTCGGCGTATACCTTAGGAG 2580
QY 2581 AGTTTTCCTCCGAACTGACAGCAACGATTCACACCGGCAAGCTGGCTTCCAACTATA 2640
Db 2581 AGTTTTCCTCCGAACTGACAGCAACGATTCACACCGGCAAGCTGGCTTCCAACTATA 2640
QY 2641 ACATTTACTCTCAAGGTCTCTTTAAATGGGCTTGTTTTACTGCTCTCTATTTAATTC 2700
Db 2641 ACATTTACTCTCAAGGTCTCTTTAAATGGGCTTGTTTTACTGCTCTCTATTTAATTC 2700
QY 2701 CCCACTATTAGCAGCTTTTGATTTGTGCTTAAGGAAATATATGCAAAAGAACCAAGACA 2760
Db 2701 CCCACTATTAGCAGCTTTTGATTTGTGCTTAAGGAAATATATGCAAAAGAACCAAGACA 2760
QY 2761 GAATGTATATGAAGAAATTTTAAATTTCTAAATTTAAATTTAAATTTAAATTTAGATCGTTACT 2820
Db 2761 GAATGTATATGAAGAAATTTTAAATTTCTAAATTTAAATTTAAATTTAAATTTAGATCGTTACT 2820
QY 2821 TGGAAATGGAGCTTAAGTCTGTGGGACAGATTAATTAATTTTCTGGCTGAAT 2880
Db 2821 TGGAAATGGAGCTTAAGTCTGTGGGACAGATTAATTAATTTTCTGGCTGAAT 2880
QY 2881 ATGTAGACTGTGTTTGAAGCTATGGGTTTATTTCTTAGAACATTTCTCTTTT 2940
Db 2881 ATGTAGACTGTGTTTGAAGCTATGGGTTTATTTCTTAGAACATTTCTCTTTT 2940
QY 2941 CTCATTATGTTACTTCTAGTGTTCACCTCTGATTAAGATTTCTTTGGTGAATAGAAA 3000
Db 2941 CTCATTATGTTACTTCTAGTGTTCACCTCTGATTAAGATTTCTTTGGTGAATAGAAA 3000
QY 3001 AAAAAAAAAAAAAAAAAAGGCGCGCG 3025
Db 3001 AAAAAAAAAAAAAAAAAAGGCGCGCG 3025
```

RESULT 4

US-09-291-839-3

; Sequence 3, Application US/09291839A

; Patent No, 6261818

; GENERAL INFORMATION:

```
; APPLICANT: Raju, Jeyaseelan
; TITLE OF INVENTION: NOVEL CARP PROTEIN AND NUCLEIC ACID MOLECULES AND USES THEREOF
; FILE REFERENCE: MNI-068
; CURRENT APPLICATION NUMBER: US/09/291,839A
; CURRENT FILING DATE: 1999-04-14
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2505
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2505)
US-09-291-839-3
```

```
Query Match 82.8%; Score 2505; DB 3; Length 2505;
Best Local Similarity 100.0%; Pred No. 0;
Matches 2505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 48 ATGGAAATTAATAATCTAGACCAACCCAAACTTTGCTGATGATGGAAGAAAAGTC 107
Db 1 ATGGAAATTAATAATCTAGACCAACCCAAACTTTGCTGATGATGGAAGAAAAGTC 60
QY 108 AGTGAATCATATGTTATCAATAGAAAGATTAGAGATGACCTGCAGATCAAGGAAAA 167
Db 61 AGTGAATCATATGTTATCAATAGAAAGATTAGAGATGACCTGCAGATCAAGGAAAA 120
QY 168 GAACTGACAGAACTTAAGGAATATATTGGCTCTGTATGAAGCCTTCAGTAAAGTCAATT 227
Db 121 GAACTGACAGAACTTAAGGAATATATTGGCTCTGTATGAAGCCTTCAGTAAAGTCAATT 180
QY 228 AATTACGCGACTGAAATGGGCTGTCTTACTTTCATTTATGTTGATTTGTGGAGGCAAG 287
Db 181 AATTACGCGACTGAAATGGGCTGTCTTACTTTCATTTATGTTGATTTGTGGAGGCAAG 240
QY 288 AAATCACATATTCGAACCTTTATGTTGAAGGGCTCCGCCCATCTCGACTGACAAAGAA 347
Db 241 AAATCACATATTCGAACCTTTATGTTGAAGGGCTCCGCCCATCTCGACTGACAAAGAA 300
QY 348 GGATTTACGCTTTGCAATTTAGCAGTTTACAGGATTAACAGGATTAATGCAAGTTGATCTCTG 407
Db 301 GGATTTACGCTTTGCAATTTAGCAGTTTACAGGATTAATGCAAGTTGATCTCTCTG 360
QY 408 CTTCACAGTGGAGCTGATATACAGCAGGTTTGAAGTACGGTGGCTCACTCCCTCCATATT 467
Db 361 CTTCACAGTGGAGCTGATATACAGCAGGTTTGAAGTACGGTGGCTCACTCCCTCCATATT 420
QY 468 GCTACAATAGCTGCCACCTAGAGGCTGTGATGTCTGTTCGACATGGAGCTAATGTC 527
Db 421 GCTACAATAGCTGCCACCTAGAGGCTGTGATGTCTGTTCGACATGGAGCTAATGTC 480
QY 528 AATATTCAAGATGACAGTGTTCCTTTGAAATTTGGTCTGTATGTAATGTAAGTGGTGAAGTGA 647
Db 541 CAGGTAACCTCGCTTCTTTTGAATTTGGTGTCTGATGTAATGTAAGTGGTGAAGTGA 600
QY 648 GATAGACCCCTCCACCTAGCATCTGAAAGAGTTCTTGAATATTGCAAACTCTTGATG 707
Db 601 GATAGACCCCTCCACCTAGCATCTGAAAGAGTTCTTGAATATTGCAAACTCTTGATG 660
QY 708 GAAGAAGCAGCAAGCAGATGTGAATGTCTCAAGATAATCAAGACCATCTCCACTCCAT 767
Db 661 GAAGAAGCAGCAAGCAGATGTGAATGTCTCAAGATAATCAAGACCATCTCCACTCCAT 720
QY 768 TTCTGTTCTCGAATTTGGACACCATGATATAGTTAAGTATCTGCTGCAAAAGTATTGGAA 827
Db 721 TTCTGTTCTCGAATTTGGACACCATGATATAGTTAAGTATCTGCTGCAAAAGTATTGGAA 780
QY 828 GTTCAACCTCATGTGTTGTTAATATCTATGGAGATACCCCTTACACCTGGCATGCTACAAT 887
```

Db 781 GTTCAACCTCATGTGTTAATATCTATGAGATACCCCTTACACTGCGCTGCTACAA 840  
Qy 888 GGCNAATTTCAAGTTGCCAGGAATATCCAAATATCAGGACAGAAAGTCTGACTAAG 947  
Db 841 GGCNAATTTCAAGTTGCCAGGAATATCCAAATATCAGGACAGAAAGTCTGACTAAG 900  
Qy 948 GAAAAATCTTCAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1007  
Db 901 GAAAAATCTTCAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960  
Qy 1008 CTAGTCAAAATTTCTTGTATGATGATGATGATGATGATGATGATGATGATGATGAT 1067  
Db 961 CTAGTCAAAATTTCTTGTATGATGATGATGATGATGATGATGATGATGATGATGAT 1020  
Qy 1068 CACACTGGATTTACACTCTGCTTCTACACAGGTCACATTCGCTTGGTTCAGTTCTTACTG 1127  
Db 1021 CACACTGGATTTACACTCTGCTTCTACACAGGTCACATTCGCTTGGTTCAGTTCTTACTG 1080  
Qy 1128 GATAATGGAGCTGATATGAATATAGTGGCTTGTATGCCAGAGGTCATAGTGGTGAAGAA 1187  
Db 1081 GATAATGGAGCTGATATGAATATAGTGGCTTGTATGCCAGAGGTCATAGTGGTGAAGAA 1140  
Qy 1188 GATGAGCAGACATGTTTGTATGGGCTTATGAAGAGGCGATGATGCCATTTGTACACTC 1247  
Db 1141 GATGAGCAGACATGTTTGTATGGGCTTATGAAGAGGCGATGATGCCATTTGTACACTC 1200  
Qy 1248 CTGAAGCAATTAAGAGACACAAAGATGAATTCGCTTGAATTAATTTCTCAGGCTCGA 1307  
Db 1201 CTGAAGCAATTAAGAGACACAAAGATGAATTCGCTTGAATTAATTTCTCAGGCTCGA 1260  
Qy 1308 GGAGATGGCTCCTATGATGCTGTCTTCACTACCCCTTGGGGAAGATTAAGAGATGACAAA 1367  
Db 1261 GGAGATGGCTCCTATGATGCTGTCTTCACTACCCCTTGGGGAAGATTAAGAGATGACAAA 1320  
Qy 1368 GAGAAGCAGATATTCCTCTCAAGAGTGGATTCCTTCAATTTCCATCTTCAGCTC 1427  
Db 1321 GAGNAGCAGATATTCCTCTCAAGAGTGGATTCCTTCAATTTCCATCTTCAGCTC 1380  
Qy 1428 TCAGAAATTCAGTTCCATGAGATATTTGGTTCAGTTCTTTTGGGAAGATATATAAGGA 1487  
Db 1381 TCAGAAATTCAGTTCCATGAGATATTTGGTTCAGTTCTTTTGGGAAGATATATAAGGA 1440  
Qy 1488 CGATGCAGAAATAAATAGTGGCTATTAAGAGTATTAAGAGTATTAAGAGTATTAAGAG 1547  
Db 1441 CGATGCAGAAATAAATAGTGGCTATTAAGAGTATTAAGAGTATTAAGAGTATTAAGAG 1500  
Qy 1548 TCAGATGTGATATGTTTTCCGAGAGGTGTCCATTTCTTCCAGCTCAATCATCCCTGC 1607  
Db 1501 TCAGATGTGATATGTTTTCCGAGAGGTGTCCATTTCTTCCAGCTCAATCATCCCTGC 1560  
Qy 1608 GTAATTCAGTTTGGTGGTCTTCTTGAATGATCCAGCCAGTTTGGCAATTCCTACTCA 1667  
Db 1561 GTAATTCAGTTTGGTGGTCTTCTTGAATGATCCAGCCAGTTTGGCAATTCCTACTCA 1620  
Qy 1668 TACATATCAGGGGTTCTCTGTTCTCCCTCTTCAAGAGAGAGATTTCTTGAATTTG 1727  
Db 1621 TACATATCAGGGGTTCTCTGTTCTCCCTCTTCAAGAGAGAGATTTCTTGAATTTG 1680  
Qy 1728 CAGTCTAAATTAATTAATTCAGTAGATTTGCCAAGGCGATGATGATGATGATGATGATG 1787  
Db 1681 CAGTCTAAATTAATTAATTCAGTAGATTTGCCAAGGCGATGATGATGATGATGATGATG 1740  
Qy 1788 ACACAGCAATTAATTAATTCAGTAGATTTGCCAAGGCGATGATGATGATGATGATGATG 1847  
Db 1741 ACACAGCAATTAATTAATTCAGTAGATTTGCCAAGGCGATGATGATGATGATGATGATG 1800  
Qy 1848 CATGCTGTGTGGCAGATTTTGGAGATCAAGATTTCTACAGTCTCTGATGATGATGATGATG 1907  
Db 1801 CATGCTGTGTGGCAGATTTTGGAGATCAAGATTTCTACAGTCTCTGATGATGATGATGATG 1860  
Qy 1908 ATGCAAAAACAACTGGGAACCTCCGTTGGATGGCTTCTGAGGTGTTTCCAGGTGCTACT 1967

Db 1861 ATGACAAAACAACTGGGAACCTCCGTTGGATGGCTCTGAGGTGTTTCCAGCAGTGCACT 1920  
Qy 1968 CGGTACACCATCAAGCAGATGTCTTACGATATGTCTGTGTCTGTGGAAATTTCTCACT 2027  
Db 1921 CGGTACACCATCAAGCAGATGTCTTACGATATGTCTGTGTCTGTGGAAATTTCTCACT 1980  
Qy 2028 GCGAAATTTCCATTCGCTCATCTCAAGCAGCGGTGGCGGAGCAGACATGGCTTACCAC 2087  
Db 1981 GCGAAATTTCCATTCGCTCATCTCAAGCAGCGGTGGCGGAGCAGACATGGCTTACCAC 2040  
Qy 2088 CACATCAGACCTCCCATTTGGCTATTCCATTTCCCAAGCCCATATCATCTCTGCTGATAGA 2147  
Db 2041 CACATCAGACCTCCCATTTGGCTATTCCATTTCCCAAGCCCATATCATCTCTGCTGATAGA 2100  
Qy 2148 GGGTGAACGCATGTCTCTGAAGGAAGACCCGAATTTCTGAAGTTGTCTGAAGTTAGAA 2207  
Db 2101 GGGTGAACGCATGTCTCTGAAGGAAGACCCGAATTTCTGAAGTTGTCTGAAGTTAGAA 2160  
Qy 2208 GAGTGTCTCTGCAACATTTGAGCTGATGTCTCTGATCAAGTAAACAGAGTGGGTCTCTC 2267  
Db 2161 GAGTGTCTCTGCAACATTTGAGCTGATGTCTCTGATCAAGTAAACAGAGTGGGTCTCTC 2220  
Qy 2268 TCACCTTCTTCTTCTTCTGATTTGCCTGTGAACCGGGGAGGACCTGGCCGGAGTCATGTG 2327  
Db 2221 TCACCTTCTTCTTCTGATTTGCCTGTGAACCGGGGAGGACCTGGCCGGAGTCATGTG 2280  
Qy 2328 GAGCATTAAGAAAGTCTGTTTGAATTTGAAATGCTCTTAAATGCAAGTCTCTATGCTGCT 2387  
Db 2281 GCAGCATTAAGAAAGTCTGTTTGAATTTGAAATGCTCTTAAATGCAAGTCTCTATGCTGCT 2340  
Qy 2388 TTGTCCCAAGTCTCGCAATATTCTCTCAAGTCTGTCTTTGGAGAGATGAAAGA 2447  
Db 2341 TTGTCCCAAGTCTCGCAATATTCTCTCAAGTCTGTCTTTGGAGAGATGAAAGA 2400  
Qy 2448 AGTCTTCAATACACCCATTCGAAATATGCTATGATGATGATGATGATGATGATGATG 2507  
Db 2401 AGTCTTCAATACACCCATTCGAAATATGCTATGATGATGATGATGATGATGATGATG 2460  
Qy 2508 CATTTTCATTTGCGGAATAGTAGCAGCTTTGAGGACGAGC 2552  
Db 2461 CATTTTCATTTGCGGAATAGTAGCAGCTTTGAGGACGAGC 2505

RESULT 5  
US-09-458-457-3  
; Sequence 3, Application US/09458457  
; Patent No. 650654  
; GENERAL INFORMATION:  
; APPLICANT: Raju, Jeyaseelan  
; TITLE OF INVENTION: NOVEL CARP PROTEIN AND NUCLEIC ACID MOLECULES AND USES  
; TITLE OF INVENTION: THEREFOR  
; FILE REFERENCE: nri-068CD  
; CURRENT APPLICATION NUMBER: US/09/458,457  
; EARLIER FILING DATE: 1999-12-10  
; EARLIER FILING DATE: 1998-12-11  
; EARLIER APPLICATION NUMBER: 09/291,839  
; EARLIER FILING DATE: 1999-04-14  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 2505  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(2505)  
; US-09-458-457-3

Query Match 82.8%; Score 2505; DB 4; Length 2505;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	48	ATGGGAATTAATAATCTAGACCAACCCAACTTGTACTGATGAATGGAAGAAAGTC	107	Db	1081	GATTAATGGAGCTGATATGAATCTAGTGGCTTGTGATCCCGCAGAGTCTAGTGGTGA	1140
Db	1	ATGGGAATTAATAATCTAGACCAACCCAACTTGTACTGATGAATGGAAGAAAGTC	60	QY	1188	GATGAGCAGACATGTTTGTATGATGGCTTATGAAAAAGGGCATGATGCCATTTGTCACTC	1247
QY	108	AGTGAATCATATGTTTATCACAAATAGAAAGATAGAAAGATGACCTGCGAGATCAAGGAAAA	167	Db	1141	GATGAGCAGACATGTTTGTATGATGGCTTATGAAAAAGGGCATGATGCCATTTGTCACTC	1200
Db	61	AGTGAATCATATGTTTATCACAAATAGAAAGATAGAAAGATGACCTGCGAGATCAAGGAAAA	120	QY	1248	CTGAAGCATTTATAGAGACCAACAGATGAATTCGCCCTGTAAATCAATATTCTCAGCTGGA	1307
QY	169	GAACTGACAGAACTTAAGGAATATATTTGGCTCTGATGAAGCCTTCAGTAAATCAATTTA	227	Db	1201	CTGAAGCATTTATAGAGACCAACAGATGAATTCGCCCTGTAAATCAATATTCTCAGCTGGA	1260
Db	121	GAACTGACAGAACTTAAGGAATATATTTGGCTCTGATGAAGCCTTCAGTAAATCAATTTA	180	QY	1308	GGAGATGGCTCTATGATGCTGTTCCATCACCCCTGGGGAAGATTAAGAGCATGACAAA	1367
QY	228	AATTACCGCACTGAAAAATGGGCTGCTCTACTTCAATTTATGTTGCAATTTGTGAGGCAAG	287	Db	1261	GGAGATGGCTCTATGATGCTGTTCCATCACCCCTGGGGAAGATTAAGAGCATGACAAA	1320
Db	181	AATTACCGCACTGAAAAATGGGCTGCTCTACTTCAATTTATGTTGCAATTTGTGAGGCAAG	240	QY	1368	GAGAAGCAGATATTCCTCCCTAAGAGCTGGATTCGCTTCAATTTCCATTTCCAGCTC	1427
QY	288	AAATCACATATTCGAACCTTATATGTTGAAAGGGCTCCGCCCATCTCGACTGACAGAAAT	347	Db	1321	GAGAAGCAGATATTCCTCCCTAAGAGCTGGATTCGCTTCAATTTCCATTTCCAGCTC	1380
Db	241	AAATCACATATTCGAACCTTATATGTTGAAAGGGCTCCGCCCATCTCGACTGACAGAAAT	300	QY	1428	TCAGAAATTCAGATTCATGAGATTAATGGCTCAGGTTCTTTTGGGAAAGTATATAAGGA	1487
QY	348	GGATTTACAGCCTTCATTTAGCAGTTTACAGGATATATGAGAAATTTGATCACTTCTCTG	407	Db	1381	TCAGAAATTCAGATTCATGAGATTAATGGCTCAGGTTCTTTTGGGAAAGTATATAAGGA	1440
Db	301	GGATTTACAGCCTTCATTTAGCAGTTTACAGGATATATGAGAAATTTGATCACTTCTCTG	360	QY	1488	CGATGCAAAATAAATAGTGGCTATAAAGCTTATCGAGCCAAATACCTACTGCTCCAG	1547
QY	408	CTTCAGTGGAGCTGATATACAGCAGGTTGGATACGGTGGCTCTCACTGCCCTCCATATT	467	Db	1441	CGATGCAAAATAAATAGTGGCTATAAAGCTTATCGAGCCAAATACCTACTGCTCCAG	1500
Db	361	CTTCAGTGGAGCTGATATACAGCAGGTTGGATACGGTGGCTCTCACTGCCCTCCATATT	420	QY	1548	TCAGATGATATATGTTTCCGAGAGGTGTCCATTTCTCTGCGAGCTCAATCATCCCTGC	1607
QY	468	GCTAATAAGTGGCCACTAGAGCTGCTGATGCTGTTGCAACATGAGCTAATGTC	527	Db	1501	TCAGATGATATATGTTTCCGAGAGGTGTCCATTTCTCTGCGAGCTCAATCATCCCTGC	1560
Db	421	GCTAATAAGTGGCCACTAGAGCTGCTGATGCTGTTGCAACATGAGCTAATGTC	480	QY	1608	GTAATTCAGTTTGTGGGTGCTTCTGTAATGATCCAGCAGTTTGCATTTGTCTACTCAA	1667
QY	528	AATATTCAAGATGCAATTTTTTTTCACTCCATTTGATATTCAGCGTACTATGACATGAA	587	Db	1561	GTAATTCAGTTTGTGGGTGCTTCTGTAATGATCCAGCAGTTTGCATTTGTCTACTCAA	1620
Db	481	AATATTCAAGATGCAATTTTTTTTCACTCCATTTGATATTCAGCGTACTATGACATGAA	540	QY	1668	TACATATCAGGGGTCTCTGTTCTCCCTCATGAGCAGAAAGAGGATTTCTGATTTG	1727
QY	589	CAGGTAATCTGGCTTTTGAATTTGGTGCTGCTGATGTAATGTAAGTGTGAAGTTGGA	647	Db	1621	TACATATCAGGGGTCTCTGTTCTCCCTCATGAGCAGAAAGAGGATTTCTGATTTG	1680
Db	541	CAGGTAATCTGGCTTTTGAATTTGGTGCTGCTGATGTAATGTAAGTGTGAAGTTGGA	600	QY	1728	CAGTCTAAATTAATATTCAGTAGATTTGCCAAAGGATGGAGTACCTTCAACACCTG	1787
QY	648	GATAGACCCCTCCACTAGCATCTGCAAAAGGATTTCTTGAATATGCAAACTCTTGAATG	707	Db	1681	CAGTCTAAATTAATATTCAGTAGATTTGCCAAAGGATGGAGTACCTTCAACACCTG	1740
Db	601	GATAGACCCCTCCACTAGCATCTGCAAAAGGATTTCTTGAATATGCAAACTCTTGAATG	660	QY	1788	ACAGAGCCATTAATCATCGTGATTTGAACAGTCAATATTTCTCTCTATCAGGATGG	1847
QY	708	GAAGAAGCAGCAAGCAGATGTGAATGCTCAAGATATGAAGACCATGTCACCTCCAT	767	Db	1741	ACAGAGCCATTAATCATCGTGATTTGAACAGTCAATATTTCTCTCTATCAGGATGG	1800
Db	661	GAAGAAGCAGCAAGCAGATGTGAATGCTCAAGATATGAAGACCATGTCACCTCCAT	720	QY	1848	CATGCTGTGTGGCAGATTTTGGAGAAATCAAGATTTTCTACAGTCTCTGGATGAAGCAAC	1907
QY	768	TTCTGTTCTCGATTTGGACACCATGATATAGTTAAGTATCTGCTGCAAGTGAATTTGGA	827	Db	1801	CATGCTGTGTGGCAGATTTTGGAGAAATCAAGATTTTCTACAGTCTCTGGATGAAGCAAC	1860
Db	721	TTCTGTTCTCGATTTGGACACCATGATATAGTTAAGTATCTGCTGCAAGTGAATTTGGA	780	QY	1908	ATGCAAAAAACCTGGAAACCTCCGTTGGATGGCTCTCAGGTTTACGAGTGCCT	1967
QY	828	GTTCACCTCATGTTGTTAATATCTATGAGATACCCCTTACACCTGGCATCTCAAT	887	Db	1861	ATGCAAAAAACCTGGAAACCTCCGTTGGATGGCTCTCAGGTTTACGAGTGCCT	1920
Db	781	GTTCACCTCATGTTGTTAATATCTATGAGATACCCCTTACACCTGGCATCTCAAT	840	QY	1968	CGGTACACCATCAAGCAGATGCTTTCAGCTATGCTCTGTCTGTGGGAAATTTCTCACT	2027
QY	888	GGCAATTTGAAGTTGCCAAGAAATCATCCAAATATCAGGAACAGAACTCTGACTAAG	947	Db	1921	CGGTACACCATCAAGCAGATGCTTTCAGCTATGCTCTGTCTGTGGGAAATTTCTCACT	1980
Db	841	GGCAATTTGAAGTTGCCAAGAAATCATCCAAATATCAGGAACAGAACTCTGACTAAG	900	QY	2028	GGCGAAATTCATTCGCTCATCTCAAGCAGCGGCTGGCGCAGCAGACATGGCTTACAC	2087
QY	948	GAACAACATCTTCAGTGAACACAGCTTTTTCATAGTCTGTTGATCTATGGCAAGCATTTGAC	1007	Db	1981	GGCGAAATTCATTCGCTCATCTCAAGCAGCGGCTGGCGCAGCAGACATGGCTTACAC	2040
Db	901	GAACAACATCTTCAGTGAACACAGCTTTTTCATAGTCTGTTGATCTATGGCAGACATTGAC	960	QY	2088	CACATCAGACTCCCATTTGGCTATTCATTTCCCAAGCCCATATCATCTCTGCTGATACGA	2147
QY	1008	CTAGTCAAAATTTCTTTCTTGATCAGAAATGTCATAAACAATCAACCAAGAGGGATGGG	1067	Db	2041	CACATCAGACTCCCATTTGGCTATTCATTTCCCAAGCCCATATCATCTCTGCTGATACGA	2100
Db	961	CTAGTCAAAATTTCTTTCTTGATCAGAAATGTCATAAACAATCAACCAAGAGGGATGGG	1020	QY	2148	GGGTGGAACCATGTCTGTAAGGAAGACCCGAAATTTTCTGAAGTTGTCTATGAAGTTAGAA	2207
QY	1068	CACACTGATTAACCTCTGCTTGTACCAAGGTCACATTCGCTGGTTTCAAGTTCTTACTG	1127	Db	2101	GGGTGGAACCATGTCTGTAAGGAAGACCCGAAATTTTCTGAAGTTGTCTATGAAGTTAGAA	2160
Db	1021	CACACTGATTAACCTCTGCTTGTACCAAGGTCACATTCGCTGGTTTCAAGTTCTTACTG	1080	QY	2208	GAGTGTCTCTGCAACATTTGATGATGCTCTCTGATCAAGTAAACAGAGCTGGTCTCTC	2267
QY	1128	GATAATGGAGCTGATATGAATCTAGTGGCTTGTGATCCCGCAGAGCTTAGTGGTGA	1187				

Db	2161	GAGTGTCTCTGCAACATTGAGCTGATGTCTCTCTGCAATCAAGTAAACAGCAGTGGGTCTCTC	2322
QY	2268	TCACCTCTCTCTCTCTCTCTGATTCGCTGGTGAACCGGGGAGGACCTGGCCGGAGTCAATGTG	2327
Db	2221	TCACCTCTCTCTCTCTCTGATTCGCTGGTGAACCGGGGAGGACCTGGCCGGAGTCAATGTG	2280
QY	2328	GCAGCATTAAGAAGTGGTTTCGAATTGGAATATGCTCTTAATGCAAGTCTCTATGCTGCT	2387
Db	2281	GCAGCATTAAGAAGTGGTTTCGAATTGGAATATGCTCTTAATGCAAGTCTCTATGCTGCT	2340
QY	2388	TTGTCCCAAAAGTGTGCACAAATATTCCTCTCAAGTCTCTCTTTGGAGGAGATCAAAAGA	2447
Db	2341	TTGTCCCAAAAGTGTGCACAAATATTCCTCTCAAGTCTCTCTTTGGAGGAGATCAAAAGA	2400
QY	2448	AGTCCTCAATACACACCAATGACAAATATGGCTATGATTCGATCCCATGAGTCAATG	2507
Db	2401	AGTCCTCAATACACACCAATGACAAATATGGCTATGATTCGATCCCATGAGTCAATG	2460
QY	2508	CATTTCCTCAATTCCTGCCAAATAGTAGCAGCTTTTGAGACAGCAGC	2552
Db	2461	CATTTCCTCAATTCCTGCCAAATAGTAGCAGCTTTTGAGACAGCAGC	2505
RESULT 6			
US-09-947-199A-3			
; Sequence 3, Application US/09947199A			
; Patent No. 6660490			
; GENERAL INFORMATION:			
; APPLICANT: Raju, Jayaseelan			
; TITLE OF INVENTION: NOVEL CARC PROTEIN AND NUCLEIC ACID MOLECULES AND USES			
; TITLE OF INVENTION: THEREFOR			
; FILE REFERENCE: MNI-068CP2			
; CURRENT APPLICATION NUMBER: US/09/947,199A			
; CURRENT FILING DATE: 2001-09-05			
; PRIOR APPLICATION NUMBER: US 60/111,938			
; PRIOR FILING DATE: 1998-12-11			
; PRIOR APPLICATION NUMBER: US 09/291,839			
; PRIOR FILING DATE: 1999-04-14			
; PRIOR APPLICATION NUMBER: US 09/458,457			
; PRIOR FILING DATE: 1999-12-10			
; NUMBER OF SEQ ID NOS: 9			
; SOFTWARE: Patent Ver. 2.0			
; SEQ ID NO 3			
; LENGTH: 2505			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
; FEATURE:			
; NAME/KEY: CDS			
; LOCATION: (1)..(2505)			
US-09-947-199A-3			
Query Match 82.8%; Score 2505; DB 4; Length 2505;			
Best Local Similarity 100.0%; Pred. No. 0;			
Matches 2505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	48	ATGGGAATTTATAATCTAGACCAACCCAACTGTACTGATGATCGAAGAAAAAGTC	107
Db	1	ATGGGAATTTATAATCTAGACCAACCCAACTGTACTGATGATCGAAGAAAAAGTC	60
QY	108	AGTGAATCATATGTTTATCACAATAGAAGATTAGAAGTCACTCGAGATCAAGGAAAAA	167
Db	61	AGTGAATCATATGTTTATCACAATAGAAGATTAGAAGTCACTCGAGATCAAGGAAAAA	120
QY	168	GAACTGACAGAACTAAGGAATATATTTGGCTCTGATGAAGCCCTCAGTAAAGTCAATTTA	227
Db	121	GAACTGACAGAACTAAGGAATATATTTGGCTCTGATGAAGCCCTCAGTAAAGTCAATTTA	180
QY	226	AATTACCGCACTGAAATGGGCTGTCTACTCATTTATTTGTCATTTTGGAGGCAAG	287
Db	181	AATTACCGCACTGAAATGGGCTGTCTACTCATTTATTTGTCATTTTGGAGGCAAG	240
QY	288	AAATCATCATTTTGAACCTTTATGTTGAAAGGGCTCGGCCATCTCGACTCAAGAAAT	347

241	AAATCACATATTGCAACTCTTATGTTGAAAGGGCTCCGCCCATCTCGACTGACACAGAAT	300
348	GGATTTACAGCCTTGCATTTAGCAGTTTACAAAGGATAATCAGAGTAATGATCACTTCTCTCG	407
301	GGATTTACAGCCTTGCATTTAGCAGTTTACAAAGGATAATGAGAAATTCATCACTTCTCTCG	360
408	CTTCACAGTGGAGCTGATATACAGCAGGTTGATACGGTGGCCTCACTCAGCCCTCCATATT	467
361	CTTCACAGTGGAGCTGATATACAGCAGGTTGATACGGTGGCCTCACTCAGCCCTCCATATT	420
468	GCTACAATAGCTGCCACCTAGAGGCTGCTGATGTCTGTTGCAACATGGAGCTAATGTC	527
421	GCTACAATAGCTGCCACCTAGAGGCTGCTGATGTCTGTTGCAACATGGAGCTAATGTC	480
528	AATATTCAAGATCAGCTTTTTCACCTCCATTTGCATATTGCGAGCGTACTATGGACATGAA	587
481	AATATTCAAGATCAGCTTTTTCACCTCCATTTGCATATTGCGAGCGTACTATGGACATGAA	540
588	CAGGTAACCTCGCCTCTTTTGAAATTTGGTGTGATGTAAATGTAAGTGGTGAAGTTGGA	647
541	CAGGTAACCTCGCCTCTTTTGAAATTTGGTGTGATGTAAATGTAAGTGGTGAAGTTGGA	600
648	GATAGACCCCTCCACCTAGCATCTGCAAAAGGATTCTTGAATATTGCAAACTCTTGATG	707
601	GATAGACCCCTCCACCTAGCATCTGCAAAAGGATTCTTGAATATTGCAAACTCTTGATG	660
708	GAAGAAGCAGCAAGCAGATGTGAATGCTCAAGATTAATGAAGACCATGTCCCACTCCAT	767
661	GAAGAAGCAGCAAGCAGATGTGAATGCTCAAGTAAATGAAGACCATGTCCCACTCCAT	720
768	TTCTGTTCTCGATTTGGACACCAATGATATAGTTTAAGTATCTGCTGCAAAAGTGAATTGGAA	827
721	TTCTGTTCTCGATTTGGACACCAATGATATAGTTTAAGTATCTGCTGCAAAAGTGAATTGGAA	780
828	GTTCAACCTCATGTGTGTTAATATCTATGAGATACCCCTTACACCTGGCATGCTACAAT	887
781	GTTCAACCTCATGTGTGTTAATATCTATGAGATACCCCTTACACCTGGCATGCTACAAT	840
888	GGCAAAATTTGAAGTTGCCAAGGAAATCATCCAAATATCAGGAACAGAAAGTCTGACTAAG	947
841	GGCAAAATTTGAAGTTGCCAAGGAAATCATCCAAATATCAGGAACAGAAAGTCTGACTAAG	900
948	GAAACATCTTCAGTGAACAGCTTTTTCATAGTGCTTGTAACCTATGCAAGAGGATTCAC	1007
901	GAAACATCTTCAGTGAACAGCTTTTTCATAGTGCTTGTAACCTATGCAAGAGGATTCAC	960
1008	CTAGTCAAAATTTCTTCTTGATCAGAATGTCAATAACATCAACACCAAGGAAGGATGGG	1067
961	CTAGTCAAAATTTCTTCTTGATCAGAATGTCAATAACATCAACCAAGGAAGGATGGG	1020
1068	CACACTGGATTACACTCTGCTTGCACAGGTCACATTCGCTCGTTTTCAGTTCTTTACTG	1127
1021	CACACTGGATTACACTCTGCTTGCACAGGTCACATTCGCTCGTTTTCAGTTCTTTACTG	1080
1128	GATAATGGAGCTGATATGAATCTAGTGGCTTTGATCCACAGCAGCTCTAGTGGTGAATAA	1187
1081	GATAATGGAGCTGATATGAATCTAGTGGCTTTGATCCACAGCAGCTCTAGTGGTGAATAA	1140
1188	GATGACGACATGTTTGAATGTGGGCTTTATGAAAAGGCAATGATGCCATTTGCACATC	1247
1141	GATGACGACATGTTTGAATGTGGGCTTTATGAAAAGGCAATGATGCCATTTGCACATC	1200
1248	CTGAGCAATTATAGACACACAGATGAATGTCCTGTAAATATTTCTCAGCCTGGA	1307
1201	CTGAGCAATTATAGACACACAGATGAATGTCCTGTAAATATTTCTCAGCCTGGA	1260
1308	GGAGATGGCTCCTATGTCTGTGTCCATCACCTTGGGGAAGATTAAAGCATGACAAAA	1367
1261	GGAGATGGCTCCTATGTGTCTGTGTCCATCACCTTGGGGAAGATTAAAGCATGACAAAA	1320
1368	GAGAGGCAGATATTTCTCTCAAGAGCTGGATGTCCTTCAATTTCCATCTTCAGCTC	1427
1321	GAGAGGCAGATATTTCTCTCAAGAGCTGGATGTCCTTCAATTTCCATCTTCAGCTC	1380

Qy	1428	TCAGAAATTGAGTTC	CCATGAGATTATT	TGGCTCAGGTTCTTTTGGGAAAGTATATAAAGGA	1487
Db	1381	TCAGAAATTGAGTTC	CCATGAGATTATT	TGGCTCAGGTTCTTTTGGGAAAGTATATAAAGGA	1440
Qy	1488	CGATGCAGAAATAA	AAATAGTGGCTATA	AAAAAGTTATTCGAGCCAAATACCTPACTGCTCCAAAG	1547
Db	1441	CGNTGCAGAAATAA	AAATAGTGGCTATA	AAAGGTTATTCGAGCCNAATACCTACTGCTCCAAAG	1500
Qy	1548	TCAGATGTGGATAT	GTGTTTGGCCGAGAGAGTG	TCOAATCTCTGCGCAGCTCAATCATCCCTGCG	1607
Db	1501	TCAGATGTGGATAT	GTGTTTGGCCGAGAGAGTG	TCOAATCTCTGCGCAGCTCAATCATCCCTGCG	1560
Qy	1608	GTAATTCAGTTTGC	GGGTGCTTGTCTGAATGATCC	CAGCCAGTTTGGCCATTGTGCACACAA	1667
Db	1561	GTAATTCAGTTTGC	GGGTGCTTGTCTGAATGATCC	CAGCCAGTTTGGCCATTGTGCACACAA	1620
Qy	1668	TACATATCAGGGGGT	CTCTCTTCTCCCTCTTTTCATG	CAGAGAAGGATTTCTTGAATTTG	1727
Db	1621	TACATATCAGGGGGT	CTCTCTTCTCCCTCTTTTCATG	CAGAGAAGGATTTCTTGAATTTG	1680
Qy	1728	CAGTCTAAATTAAT	TATTCGAGTAGATGTTGCC	CAAGGCATGGAGTACTCTCACAACCTG	1787
Db	1681	CAGTCTAAATTAAT	TATTCGAGTAGATGTTGCC	CAAGGCATGGAGTACTCTCACAACCTG	1740
Qy	1788	ACACAGCCCAATT	ATATCATCTGAGCTTGAAC	CAGTCACAATATTTCTCTATGAGGATGGG	1847
Db	1741	ACACAGCCCAATT	ATATCATCTGAGCTTGAAC	CAGTCACAATATTTCTCTATGAGGATGGG	1800
Qy	1848	CATGCTGTGGTGG	CAGATTTTGGAGATCAAGAT	TTCTACGTCTCTGGATGAGACAAC	1907
Db	1801	CATGCTGTGGTGG	CAGATTTTGGAGATCAAGAT	TTCTACGTCTCTGGATGAGACAAC	1860
Qy	1908	ATGACAAAAACCT	GGGAAACCTCCGTTGGATGGCT	CCTGAGGTGTTCAAGGATGCACT	1967
Db	1861	ATGACAAAAACCT	GGGAAACCTCCGTTGGATGGCT	CCTGAGGTGTTCAAGGATGCACT	1920
Qy	1968	CGGTACACATCA	AAAGCAGATGTCTTCAGCTAT	GTCTGTGTCTGTGGGAAATTTCTCACT	2027
Db	1921	CGGTACACATCA	AAAGCAGATGTCTTCAGCTAT	GTCTGTGTCTGTGGGAAATTTCTCACT	1980
Qy	2028	GGCGAAATTCAT	TTCGCTCATCTCAAGCAGCGGCT	CGCGCAGCAGACATGGCTTACCAC	2087
Db	1981	GGCGAAATTCAT	TTCGCTCATCTCAAGCAGCGGCT	CGCGCAGCAGACATGGCTTACCAC	2040
Qy	2088	CACATCAGACCT	CCCATTTGGCTATTCCATTTCC	CAAGCCCATATCTCTCTGCTGATACGA	2147
Db	2041	CACATCAGACCT	CCCATTTGGCTATTCCATTTCC	CAAGCCCATATCTCTCTGCTGATACGA	2100
Qy	2148	GGGTGAACGCA	TGCTTCGAAGGAAGACCGAA	TTTTTCTGAAGTTGTCTATGAAGTTAGAA	2207
Db	2101	GGGTGAACGCA	TGCTTCGAAGGAAGACCGAA	TTTTTCTGAAGTTGTCTATGAAGTTAGAA	2160
Qy	2208	GAGTGTCTCTG	CAACATTGAGCTGATGTCTCTG	CAATCAAGTAAACAGCAGTGGGTCTCTC	2267
Db	2161	GAGTGTCTCTG	CAACATTGAGCTGATGTCTCTG	CAATCAAGTAAACAGCAGTGGGTCTCTC	2220
Qy	2268	TCACCTCTCTCT	CTCTCTGATGCTGGTGAAACCG	GGGAGGACCTGGCGGAGTCAATGTG	2327
Db	2221	TCACCTCTCTCT	CTCTCTGATGCTGGTGAAACCG	GGGAGGACCTGGCGGAGTCAATGTG	2280
Qy	2328	GCAGCATTAA	GAAAGTCGTTTCGAATTGGAA	TATGTCTCTAAATGCAAGTCTCTATGCTGCT	2387
Db	2281	GCAGCATTAA	GAAAGTCGTTTCGAATTGGAA	TATGTCTCTAAATGCAAGTCTCTATGCTGCT	2340
Qy	2388	TTGTCCCAAGT	GCTGGACAATATTCCTCTC	AGAGTCTCTTTGAGGAGATCAAAAGA	2447
Db	2341	TTGTCCCAAGT	GCTGGACAATATTCCTCTC	AGAGTCTCTTTGAGGAGATCAAAAGA	2400
Qy	2448	AGTCTTCAAT	ACACACCCANTGACAAATAT	GGCTATGTATCCGATCCCATGAGTCAATG	2507
Db	2401	AGTCTTCAAT	ACACACCCANTGACAAATAT	GGCTATGTATCCGATCCCATGAGTCAATG	2460

QY	2508	CAATTTTCATTCTTGGCCGAATAGTAGCAGCTTTGAGGACACAGCAGC	2552
DB	2461	CATTTTCATTCTTGGCCGAATAGTAGCAGCTTTGAGGACACAGCAGC	2505
RESULT 7			
US-09-458-457-7			
; Sequence 7, Application US/09458457			
; Patent No. 6500654			
; GENERAL INFORMATION:			
; APPLICANT: Raju, Jeyaseelan			
; TITLE OF INVENTION: NOVEL CARP PROTEIN AND NUCLEIC ACID MOLECULES AND USES			
; TITLE OF INVENTION: THEREFOR			
; FILE REFERENCE: mai-068cp			
; CURRENT APPLICATION NUMBER: US/09/458,457			
; CURRENT FILING DATE: 1999-12-10			
; EARLIER APPLICATION NUMBER: 60/111,938			
; EARLIER FILING DATE: 1998-12-11			
; EARLIER APPLICATION NUMBER: 09/291,839			
; EARLIER FILING DATE: 1999-04-14			
; NUMBER OF SEQ ID NOS: 9			
; SOFTWARE: PatentIn ver. 2.0			
; SEQ ID NO 7			
; LENGTH: 3026			
; TYPE: DNA			
; ORGANISM: Rattus norvegicus			
; FEATURE:			
; NAME/KEY: CDS			
; LOCATION: (61)..(2565)			
US-09-458-457-7			
Query Match 66.4%; Score 2010; DB 4; Length 3026;			
Best Local Similarity 80.6%; Pred. No. 0;			
Matches 2425; Conservative 0; Mismatches 570; Indels 13; Gaps 6;			
QY	16	CGGCCCTGGAGAAAGGAAGAACTTATATAAATGGGAATTTAAATCTAGACCAACCC	75
DB	29	CAGCACAGGAGAAAAGCAAAAGCTTCTTTAAAAATGGGAATTTACAAATCCAGACCAACAC	88
QY	76	AAACTTGTACTGATGATGGAAGAAAAGTCAAGTCAATCATATGTTTACAAATAGAAA	135
DB	89	AGACTTGTTCTGATGAATGGAAGAGAAAGTTAGTTGAATCTTACGCTATTATCATAGAAA	148
QY	136	GATTAGAAGATGACCTGCGAGATCAAGGAAAAAGAACTGCAGAACTAAGGAATATATTG	195
DB	149	GGCTGGAGGATAACCTGCGAGATCAAGAAAAATGAAATTCAGAAACTAAGSCACATCTTTG	208
QY	196	GCTCTGATGAAGCCTTCAGTAAAGTCAATTTAAATACCGCACTGAAATGGCTGTCTC	255
DB	209	GCTCTGATGAAGCCTTCAGTAAAGTCAAGTAAATTTACCGCACAGACGCTGGCCTGTCC	268
QY	256	TACTTCATTTATGTTGCTATTGTTGGAGGCAAGAAATCACATATTCGAACTCTTATGTGA	315
DB	269	TGCTACACTCTGCTGTGCTGTGGCGGCAACAGTCACATATCCGTGCCCTATGTTAA	328
QY	316	AAGGGCTCCGCCCATCTCGACTGCAAGAAATGGATTTACAGCCTTGCAATTTAGCAGTTT	375
DB	329	AAGGGCTCCGCTCCATCCAGACTCACGAGAAATGGGTTTCAGCTCTGCACTGGCCGTTT	388
QY	376	ACAGGATATGAGAAATGATCACTTCTCTGCTTCCAGTGGAGCTGATATACAGCAGG	435
DB	389	ACAAGGACGCCCGGAACCTTATCACTTCACTGTTGCACAGCGAGAGATGTTCAACAAG	448
QY	436	TTGGATACGGTGGCCTCACTGCCCTCCATATTGCTACAAATAGCTGGCCACCTAGAGGCTG	495
DB	449	TGGATACGGTGGCCTCACAGCCCTCCACATAGCTGCAATAGCTGGACACCCAGAGGCTG	508
QY	496	CTGATGCTGTTGGCAACATGGAGCTTAATGTCATATTCAGATGACAGTTTTTTTCACTC	555
DB	509	CAGAAGTGCTGCTACAAACATGGGGCCAATGTGAATGTTCAAGATGCCGCTTCTTCCACC	568
QY	556	CATTGTCATATTGACGGGTACTATGGACATGAAACAGGTAACCTGCGCTTCTTTTGAATTTG	615

Db 569 CACTGCACATTGACGCTACTATGGGACGACGAGGTAAACGAGTGCCTTTTGAAGTTTG 628  
Qy 616 GTGCTGATGTAATGTAAGTGTGAAGTTGGAGATAGA CCCCCTCCACCTAGCAATCTGCAA 675  
Db 629 GTGCTGATGTAATGTAAGCGGTGAAGTTGGGACAGGCTCTGCACTCGGCTCTGCAA 688  
Qy 676 AAGGATTTCTTGAATATTGCAAACTCTTGATGGAAGGACGAGCAAGGATGTGAATG 735  
Db 689 AGGCTTCTTCAACATTTGTGAACCTCTGTTAGAAAGGAGGCAAGGAGATGTGAAG 748  
Qy 736 CTCGAATATGAAAGACCATGTCCCATCCATTTCTGTTCTGATTTGGAACACCATGATA 795  
Db 749 CTCAGGACAAATGAAGACCACTCCCTCTGCACTTCTGTTCTGATTTGACACCAATA 808  
Qy 796 TAGTTAGTATCTGCTGCAAGTGAATTTGGAAGTTCAACCTCATGTTGTTATATCTATG 855  
Db 809 TAGTGACCTACCTGCTCCAGAGTGACTTAGAGTCCAGCTCAGTCAITTAACATCTATG 868  
Qy 856 GAGATACCCCTTACACCTGGCATGTCTACAAATGGCAAAATTTGAAGTTGCAAGGAAATCA 915  
Db 869 GTGACACTCTCTTGGCACTCGCATGCTTACAATGGAAATTTTGAAGTTGCAAGGAAATG 928  
Qy 916 TCCAAATATCAGGAACAGAAAGTCTGACTAAGGAAACATCTTCAGTGAACAGCTTTTC 975  
Db 929 TCCAGGTAACAGGAACCTGAAAGTCTGACTAAGGAAACATCTTCAGCGAGACAGCTTTTC 988  
Qy 976 ATAGTGCTTGTAACCTATGGCAAGCAATTCACCTAGTCAAAATTTCTTTGATCAGAATG 1035  
Db 989 ACAGTGCTTGTAACCTATGGCAAGCAATTCACCTAGTCAAAATTTCTTTGATCAGAATG 1048  
Qy 1036 TCATAACATCAACCAAGGAAGGATGGGCACTGGAATACACTCTGTTGCTAC 1095  
Db 1049 CTGTGAACATTAACCAAGGAAGGATGGGCACTGGAATTCAGTCTGTTGCTAC 1108  
Qy 1096 ACAGTCACTATCGCTCGTTCACTTCTTCTGATTAATGAGCTGATATGAATCTAGTGG 1155  
Db 1109 ACAGGCAATATCGGCTCGTTCACTTCTTCTGATTAATGAGCTGATATGAATCTAGTGG 1168  
Qy 1156 CTGTGATCCAGCAGCTATGAGTGAAGGATGAGCAGACATGTTTGAATGCTGGCTT 1215  
Db 1169 CTGTGATCCAGCAGCTATGAGTGAAGGATGAGCAGACATGTTTGAATGCTGGCTT 1228  
Qy 1216 ATGAAAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1275  
Db 1229 ACAGAAAGGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1288  
Qy 1276 AATGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1335  
Db 1289 AGCTGCCATGTAACGAATATCCAGCTGAGGAGATGCTCTCTATGTTCTGTTCTTT 1348  
Qy 1336 CACCTTTGGGAGATTAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1395  
Db 1349 CCCCCTTTGGGAGATTAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1408  
Qy 1396 CTGATTTGCTTCACTTCCAGCTCTCAGAAATGAGTTGATTTCCATGATGATTTG 1455  
Db 1409 CTGAATACCTTCCGCTTCCATCTCCAGCTCTCCAGCTCTCCAGCTCTCCAGCTCTCCAG 1468  
Qy 1456 GCTCAGGTTCTTTTGGGAAAGTATATAAGGAGGATGATGATGATGATGATGATGATGATGAT 1515  
Db 1469 GCTCGGTTCTTTTGGGAAAGTATATAAGGAGGATGATGATGATGATGATGATGATGATGATG 1528  
Qy 1516 AACGTTATCGGCAATACCTGCTCCAGCTCAGATGATGATGATGATGATGATGATGATGATG 1575  
Db 1529 AACGATACCGAGCAACACTGCTCCAGCTCAGATGATGATGATGATGATGATGATGATGATG 1588  
Qy 1576 TGTCCATTTCTGCGAGCTCAATCATCCCTCGCTAATTCAGTTTGTGGGTGCTGCTGATG 1635  
Db 1589 TGTCCATTTCTGCGAGCTCAACCAACCTCGCTGATGATGATGATGATGATGATGATGATGATG 1648  
Qy 1636 ATGATCCAGGAGGATTTGCAATGCTCAATACATATCAGGAGGTTTCTGTTCTTCTTCTTCT 1695  
Db 1649 ATGACCCGAGTCAAGTTTGGCAATTTGCTACTCAGTACATTTTCAGGAGGCTCCCTGTTCTCC 1708

Qy 1596 TCCTTCATGACGAGAGAGGATTTCTGATTTGAGTCTCTAAATTAATTTTTCAGTAGATG 1755  
Db 1709 TGCTTCATGAACAGAGAGAGATTTCTGATTTGAGTCTCTAAATTAATTTTTCAGTAGAG 1768  
Qy 1756 TTGCCAAAGCATGAGGATACCTTTCAACCTGACAGCAATTTATACATCTGATCTTGA 1815  
Db 1769 TTGCCAAGGSCATGAGGATACCTTGACAGCTTGACCCAGCAATCATACACCGGACCTGA 1828  
Qy 1816 ACAGTCAACATATTTCTTCTATGAGGATGGGATGCTGTGGTGGCAGATTTTGGAGAT 1875  
Db 1829 ACAGCCACAATATTTCTGCTATGAGGATGGGATGCTGTGTGGCAGATTTTGGAGAT 1888  
Qy 1876 CAAGATTTCTACAGTCTCTGGATGAAGACAAATGACAAAAACAACTGGGAACTCCCGTT 1935  
Db 1889 CAAGATTTCTGAGTCCCTGGATGAAGACAAATGACAAAGCAGCAGGAACTCGGCT 1948  
Qy 1936 GATGAGCTCTGAGGTTGTCACGAGTGCACTCGGTACACCATCAAGAGCAGATGTCTTCA 1995  
Db 1949 GGATGGCCCTGAGGTTGTCACACAGTGACGAGATACACCATCAAGGCTGATGTCTTCA 2008  
Qy 1996 GCTATGCTCTGCTGTGGGAAATTTCTCACTGGGAAATTTCCATTGCTCATCTCAAGC 2055  
Db 2009 GTTACTCTCTGCTGTGGAGCTCTCACTGGGAAATTTCCATTGCTCATCTCAAGC 2068  
Qy 2056 CAGCGCTGCGGACGACATGGCTTACCACCATCAGACCTCCCATTTGGCTATTCCA 2115  
Db 2069 CAGCGCTGCGGACGACATATGGCTATACACCATCAGACCGGCCCATCGGCTATTCCA 2128  
Qy 2116 TTCCCAAGCCATATCATCTCTGCTGATGAGGAGTGGAAACCATGCTCTGAAAGAC 2175  
Db 2129 TCCCAAGCCATCTCATCCCTGCTGATCGGGGCTGGAAATGATGCTCTGAAAGACGAC 2188  
Qy 2176 CCGAATTTCTGAAGTTGTGATGAAGTTAGAAGAGTGTCTCTGCAACATTTGAGCTGATGT 2235  
Db 2189 CAGATTTCTGAGTCTGTAGCAAACTGAGGAGTGCCTATGCAATGTGGAGCTCATGT 2248  
Qy 2236 CTCCTGATCAAGTAAACAGAGTGGTCTCTCTACCTTCTTCTTCTCTGATGCTGG 2295  
Db 2249 CTCAGCATCAAGTAAACAGAGTGGTCTCTCTGACCTTCTCTTCTCCGATTCCTGCTG 2308  
Qy 2296 TGAACCGGAGGACCTGGCGGAGTCATGTGCGAGCATTAAGAAGTCTGTTTCAATTTGG 2355  
Db 2309 TGAGCCGGGAGGAGGCTGGCGGAGCCAGTGGGAGCTTACGGAGCCTTTTGGATTGG 2368  
Qy 2356 AATATGCTCTAAATGCAAGTCTATGCTGTTTTCCTGCTGCTGCTGCTGCTGCTGCTG 2415  
Db 2369 AGTATGCCCTTAAATGCAAGTCTCTATGCTGGGTGCTCCAAAGTGTGGACACACTTA 2428  
Qy 2416 CTCAGGCTCTCTTTGGAGGAGATGAABAGAGTCTTCAATACACACCCATTGACAAAT 2475  
Db 2429 ATCCGGGCTCTCTTTGGAGGAGATGAATAGGAGCACCAGTATTCACTGTTGACAAAT 2488  
Qy 2476 ATGGCTATGATCCGATCCCATGAGCTCAATGCTTTTCTTCTTCCGAAATAGTAGCA 2535  
Db 2489 ACGGCTATGCTCTGATCCCATGAGCTGAGCAGCTTCACTCCCGCAAGACGACAGCA 2548  
Qy 2536 GCTTTGAGGACAGAGTGTGACAGCATTTGGGCTGATACCTTAAGGAGATTTTTCCTG 2595  
Db 2549 ACTTTGAGGACAGCAACTGACAG-GTCTGGCATACACCTTAAGGGGCTCTCCCATCAG 2607  
Qy 2596 CTGACGCAACGATTTCAACCCAGGCAAGCTGGCTTCCCACTATATAACATTTTACTCTCAA 2655  
Db 2608 CTGACGAGCTGATTTTACCCATGGCAGGCTTGTCTTCAATTTATACGCCCTGCTCTG 2667  
Qy 2656 AGGCTCTTAAATTTGGGCTTGTGTTTACCTGCTCTTAAATTTTCCCACTATTAGCAGG 2715  
Db 2668 AGGT-ITCTTTCAAACTGCTTGTCTTATTCTAAGCTGCTTTTAAATTTCTCTCAGGACAG 2726  
Qy 2716 CTTTGGATTTGCTCCCTAAGGAATAATATGCAAAAGAACCAAGACAGAAATGTATGAAGA 2775  
Db 2727 GCTTTGACTCATGCC--AGCCTGAAGTGTCAAGAGAGAGATACAGAAATGTGCATGAGGA 2784

QY 2776 ATTGTTTAAATTTGTAATTTAAATTTAAATTTAGATCGTTACTTGGAAATGGAGCCTA 2835  
Db 2785 ATTGTTCTTAGTTTGAATTTAAAGCCCTTAATTCCTGGGGCTGGGGTTCAAA----TC 2840  
QY 2836 AGTCTGTGGTGACAGATAAATAATTTATCTTTTCCTGGGCTGAATATGTAGACTGTGTT 2895  
Db 2841 TGTGTAGATAGCTGGGTTGACCCCTATGTATTTGTAGACCAAACTGTGTGGGCTTGTGT 2900  
QY 2896 TGACAG-CTATCGGTTTATTTCTTAGACATTTCTATTTCTTTCTCATTTATTTACT 2954  
Db 2901 TGAGGCTCTCTGTTGGGTTCTTAAACCAAGCTGGCTGAATTTATCTCTGTTGCCCTT 2960  
QY 2955 TCTAGTGTTCACCTCTGTGATTAAAGATTTCTTGGTGAATAGAAAAAATAGAAAAA 3014  
Db 2961 ----GTGTGTTACTTCTGTGATTAAAGTCTCTCGGTGATCTAGAAAAAATAGAAAAA 3016  
QY 3015 AGGGCGGC 3022  
Db 3017 GGGCGGCC 3024

## RESULT 8

US-09-947-199A-7  
; Sequence 7, Application US/09947199A  
; Patent No. 6660490

; GENERAL INFORMATION:  
; APPLICANT: Raju, Jayaseelan

; TITLE OF INVENTION: NOVEL CARP PROTEIN AND NUCLEIC ACID MOLECULES AND USES  
; TITLE OF INVENTION: THEREFOR  
; FILE REFERENCE: MNI-068CP2

; CURRENT APPLICATION NUMBER: US/09/947,199A  
; CURRENT FILING DATE: 2001-09-05

; PRIOR APPLICATION NUMBER: US 60/111,938  
; PRIOR FILING DATE: 1998-12-11

; PRIOR APPLICATION NUMBER: US 09/291,839

; PRIOR FILING DATE: 1998-04-14

; PRIOR APPLICATION NUMBER: US 09/458,457  
; PRIOR FILING DATE: 1999-12-10

; NUMBER OF SEQ ID NOS: 9

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 7

; LENGTH: 3026

; TYPE: DNA

; ORGANISM: Rattus norvegicus

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (61)..(2565)

US-09-947-199A-7

Query Match 66.4%; Score 2010; DB 4; Length 3026;  
Best Local Similarity 80.6%; Pred. No. 0;  
Matches 2425; Conservative 0; Mismatches 570; Indels 13; Gaps 6;

QY 16 CGGCCCTGGAGAAAGGAAGAACTTTATAAATGGGAAATTAATACTAGACCAACCC 75  
Db 29 CAGCAGAGAGAAAGCAAGACCTTTTAAATGGGAAATTAATAATCCAGACCAACAC 88  
QY 76 AAACCTTGACTGATGAATGGAAGAAAGTCAAGTGAATCATATGTTATCAATAGAA 135  
Db 89 AGACTGTGTTGATGAATGGAAGAAAGTGTAGTGAATCTTAGCTATTATCATAGAA 148  
QY 136 GATTAGAGATGACCTCGAGATCAAGGAAAGAAAGTCAAGAACTAAGGAATATATTG 195  
Db 149 GGCTGGAGGATAAATCGAGATCAAGAAATGAATTTCAAGAACTAAGGACATCTTTG 208  
QY 196 GCTCTGATGAACCTTCTAGTAAAGTCAATTTAAATTAACCGACATGAATAATGGGCTGTC 255  
Db 209 GCTCTGATGAACCTTCTAGTGAAGTCAAGTAAATTAACCGACAGAGCGTGGCTGTCC 268  
QY 256 TACTTCAATTTATGTTGCAATTTGTGGAGGCAAGAAATCACATATTCGAATCTTTATGTTGA 315  
Db 269 TGCTACACCTCTGCTGTCTGTGGCGGACAAAGTCAATATCCGTGGCCCTTATGTTAA 328

QY 316 AAGGCTCCGCCATCTCGACTGACAGAAATGGATTTACAGCCTTGCATTTAGCAGATTT 375  
Db 329 AAGGCTCCGCTCCATCCAGACTGACGAGAAATGGGTTCCAGCTCTGCACCTGGCGGTTT 388  
QY 376 ACAAGGATAATGACAGAAATGATCATCTTCTCTGCTTCCAGTGGAGCTGATATACAGCAG 435  
Db 389 ACAAGGACAGCCCGGAACCTTATCACTTCACTGTGTGACAGCGGAGAGATGTTACAGCAAG 448  
QY 436 TTGGATACGGTGGGCTCACCTGCCCTCCATATTCCTAATAGCTGCCACCTAGAGGCTG 495  
Db 449 TGGGATACGGTGGGCTCACAGCCTCCATAGCTGCAATAGCTGACACCCAGAGGCTG 508  
QY 496 CTGATGTCTGTTGCCAATGAGAGCTAATGTCAATATTCAAGATGAGTTTTTTTCACTC 555  
Db 509 CAGAAGTGTCTGCTACAACTGGGGCCAAATGTGAATGTTCAAGATGCGCTCTCTTCCACCC 568  
QY 556 CATTCATATTGGACGCTACTATGGACATGAACAGGTAACTGCCTCTTTTGAATTTG 615  
Db 569 CACTGCACATTGCGAGCTTACTATGGGCACGAGCAGGTAAACAGTGTCTCTTTGAAGTTG 628  
QY 616 GTGCTGATGTAATGTAAAGTGTGAAGTTGGAGATAGACCCCTCCACCTAGCATCTGCAA 675  
Db 629 GTGCTGATGTAATGTAAAGCGTGAAGTTGGGACAGGCTCTGCACTGGCTCTGCAA 688  
QY 676 AAGGATCTTGAATATTGCAAACTCTTGTATGGAAGAGGACCAAGCAGATGTAATG 735  
Db 689 AGGCTCTTCTCAACATTTGAAAACCTCTGTGTAGAAAGGAGCAAGCAGATGTGAAG 748  
QY 736 CTCGAAGATAATGAAGACCATGTCCCACTCCATTTCTGTTTCTCGATTGGACACCATGATA 795  
Db 749 CTCAGGACATGAAGACCACTGCTCTGCACTCTGTTCTCGATTGGACACCAATA 808  
QY 796 TAGTTAAGTATCTGCTGCAAAAGTGAATTTGGAAGTTCAACTATGTTGTAATATCTATG 855  
Db 809 TAGTGAGTACTCTGCTCCAGAGTGAATTTAGAGTCCAGGCTCACGTCATTAACATCTATG 868  
QY 856 GAGATACCCCTTACACCTCGCATGTACAAATGGCAAAATTTGAAGTTGCCAAGGAATCA 915  
Db 869 GTGACATCTCTTTCACCTCGCATGTACAAATGGAAATTTGAAAGTTGCCAAGGAATG 928  
QY 916 TCCAAATATCAGGAACAGAAAGTCTGACTAAGGAAACATCTTTCAGTGAAGAACAGCTTTTC 975  
Db 929 TCCAGGTAACAGGAACTGAAAGTCTGACTAAGGAAACATCTTTCAGCGAGACAGCTTTTC 988  
QY 976 ATAGTCTTGTACCTTGGCAAGACATTCACCTAGTCTCAAAATTTCTTCTGATCAGAATG 1035  
Db 989 AAGTGTGTGACCTTATGGCAAGAACTTACCTGGTCAAAATTTCTTGTATCAGAATG 1048  
QY 1036 TCATAAACATCAACCAAGGAGGATGGGCACACTGATTTACACTCTGCTTGTCTACC 1095  
Db 1049 CTGTGAACTTAACCAAGGAGAGATGGGCACACAGATTTGCACTCTGCTTGTCTACC 1108  
QY 1096 ACGTCTACATTCGCTGCTGCTTCACTTCTGATTAATGGAGCTGATTAATCTAGTGG 1155  
Db 1109 ACGGCAATATCCGCTGCTTCACTTCTGATTAATGGTGAATGATTAATCTGCTG 1168  
QY 1156 CTCTGTATCCAGCAGCTCTAGTGTGTAAGAAATGAGCAGACATGTTTGTGTTGGGCTT 1215  
Db 1169 CTCTGTATCCAGCAGCTCTAGTGTGTAAGAAATGAGCAGACATGTTTGTGTTGGGCTT 1228  
QY 1216 ATGAAAAGGCGATGATGCTTGTCTCACTCTCTGAAGCATTTATAGACACCAAGATG 1275  
Db 1229 ACAGAAAGGACATGATGCTTGTGCTTACACTCTCTGAAGCACTACAGAGACCCAGGAG 1288  
QY 1276 AATTGCCCTGTAAATGATATTTCTCAGCTGGAGAGATGGCTCTCTATGTTCTGTTCCAT 1335  
Db 1289 AGCTGCCATGTAAACGATATTTCCAGCTGGAGAGATGGCTCTCTATGTTCTGTTCTT 1348  
QY 1336 CACCTTTGGGAAAGATTAAGCATGACAAAAGAGAGGACAGATATTCCTCTCTTAAGAG 1395  
Db 1349 CCCCCCTTTGGGCAAGATTAAGCATGACAAAAGAGAGGAGATGTTCTCTCTCTGAGGG 1408  
QY 1396 CTGATTTGCCTTCCATTTCCATCTTCAGCTCTCAGAAATTTGAGTTCCATGAGATTTATG 1455

Db 1409 CTGAACCTACCTCCGCTTCCATCTCCAACTCTCGAAATCGAGTTCCACGAGATTATCG 1468  
QY GCTCAGGTTCTTTGGGAAAGTATATAGGAGCATGCGAATATAAATAGTGGCTATAA 1515  
Db 1469 GCTCGGTTCTTTGGGAAAGTCTATAAGGGGATGCGAATAAATAAATAGTGGCGATCA 1528  
QY AACGTTATCGAGCAATACCTACTGTCTCAAGTCAGATGCGATATGTTTTGCCGAGAGG 1575  
Db 1529 AACGATACCGAGCAACACACTACTGTCTCAAGTCAGATGCGATATGTTTTGCCGAGAGG 1588  
QY TGTCATTTCTGCGAGCTCAATCATCTCCCTGCGTAATTCAGTTTGGGTGCTTGCCTTGA 1635  
Db 1589 TGTCATTTCTGCGAGCTCAACACCTCCCTGCGTGTTCAGTTTGGGTGCTTGCCTTGG 1648  
QY ATGATCCAGCCAGTTTGCATTTGCTCACTCAATACATATACAGGGGTTCTCTGTCTCC 1695  
Db 1649 ATGACCCAGCTCAGTTTGCATTTGCTCACACAGTCAATTCAGGAGGCTCCCTGTTCTCC 1708  
QY TCCTTCATGAGCAGAGAGGATTTCTGATTTGAGTCTAAATTAATTAATGAGTAGATG 1755  
Db 1709 TGCTTCAATGACAGAGAGGATTTCTGATTTGAGTCTAAATTAATTAATCAATTCGGTAGAGC 1768  
QY TTGCAAGGATGAGTACTTCAAGCTGACAGCTGACAGCAATTAATACATCTGACTTGA 1815  
Db 1769 TTGCAAGGATGAGTACTTCAAGCTGACAGCTGACAGCAATTAATACATCTGACTTGA 1828  
QY ACAGTCAATATTTCTTCTATGAGATGGGATGCTGCTGGTGGCAGATTTTGAGAGAT 1875  
Db 1829 ACAGCCAAATATTTCTGCTATGAGATGGGATGCTGCTGGTGGCAGATTTTGAGAGAT 1888  
QY CAGATTTCTACAGTCTCGATGAGACAGACATGACAAACAACTGGGACCTCCGTT 1935  
Db 1889 CAGATTTCTGCACTGCTCGATGAGACAGACATGACAAACAACTGGGACCTCCGCT 1948  
QY GGATGCTCTCGAGGTGTTTCAAGCAGTGCATCGGTACACCACTCAAGCAGATGCTTCA 1995  
Db 1949 GGATGCTCTCGAGGTGTTTCAAGCAGTGCATCGGTACACCACTCAAGCAGTGCATCG 2008  
QY GCTATGCTGCTGTGTGGGAAATTTCTCATGGGGAATTCATTCGCTCATCTCAAGC 2055  
Db 2009 GTTACTCCCTGTGTGTGGGAGTCTCTCATGAGGAAATTCATTCGCTCATCTCAAGC 2068  
QY CAGCGCTGGGAGAGACATGCTTACCACCATGAGCTCCCTGCTGAGTCCCTGCTATTCCA 2115  
Db 2116 TTCCCAAGCCCATATCATCTCTGATAGAGGGTGGAAACGATGCTCTGAGGAGAC 2175  
Db 2129 TCCCAAGCCCATATCATCTCTGATAGAGGGTGGAAACGATGCTCTGAGGAGAC 2188  
QY CCGAATTTCTGAGTGTTCATGAAGTTAGAGAGTGTCTCTGCAACATTCAGCTGATGT 2235  
Db 2189 CAGAGTTCTCTGAAGTGTGTAGCAAACTGGAGAGTGCCTATGCAATGGAGCTCATGT 2248  
QY CTCCTGATCAAGTAAACAGCAGTGGGTCTCTCTCACTTCTCTCTCTGATTCGCTGG 2295  
Db 2249 CTCAGCATCAAGTAAACAGCAGTGGGTCTCTCTCACTTCTCTCTCTGATTCGCTGG 2308  
QY TGAACCGGGAGAGACCTGGCCGAGTCTATGTCAGCATTAAGAGTCTTCGAAATGG 2355  
Db 2309 TGAGCCGGGAGGCGCTGGCCGAGTCTATGTCAGCATTAAGAGTCTTCGAAATGG 2368  
QY AATATGCTCTAAATGCAAGTCTCTATGCTCTTGTGCAAAAGTGTGGCAATATCTCT 2415  
Db 2369 AGTATGCCCTTAATGCAAGTCTCTATGCTGGTGGTGGTGGTGGTGGTGGTGGTGGT 2428  
QY CTCAGGCTGTCTTTGGAGAGATGAAAGAGTCTTCAATACACACCACTGCAAAAT 2475  
Db 2429 ATCCGGCTGTCTTTGGAGAGATGAAATAGGAGACCCAGTATTCAGTGTGCAAAAT 2488  
QY ATGGCTATGATCCGATCCCATGAGCTCAATGCTATTTCTCTGCTGGCAATATGATCA 2535

Db 2489 ACGCTATGTCTGATCCCATGAGCTGACGACCTTCACTCCCGCCCAAGACGACGCA 2548  
QY GCTTTGAGGACAGCAGCTGACAGCATTCGCGGTATACCTTAAGGAGAGTCTTTTCCCGAA 2595  
Db 2549 ACITTTGAGGACAGCAACTGACAG-GTCTGGCATACACCTTAAGGGGCTCTCCCATCAGG 2607  
QY CTGACGCAACGATTCACCAACGCGCAAGCTGGCTTCCAACTATAACATTTTACTCTCAA 2655  
Db 2608 CTGACGCAAGTATTTTCCCATGGCAGGCTTCTTCCAAATATAACGCGCTGCGCTCTG 2667  
QY AGGCTCTCTTAATTTGGGCTTGTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2715  
Db 2668 AGGT-TTCTTCAATCGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2726  
QY CTITGGATTTTGGCTTAAGGAATAATATATCAAAAGAACCAAGACAGAAATGTATGAAGA 2775  
Db 2727 GCTTTGACTCATGCC--AAGCCTGAAGTGTCAAGAGCAGATACAGAAATGTGCATGAGA 2784  
QY ATTGTTTCTTAAATTTGTAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2835  
Db 2785 ATTGTTCTTAGTTTGAATTTTAAAGCCCTTAATTCCTGGGCTGGGTTCAA---TC 2840  
QY AGTCTGTGGTGGACAGATAATATATGTTTCTGGGCTGAATATGTAGACTTGTGT 2895  
Db 2841 TGTGTAGATAGCTGGGTGACCCCTTATGTATTTGTAGACCAACTGTGTGGCTTGTGT 2900  
QY TGACAG-CTATGGGTTTATTTCTTAGAACATTTGTCATTTTCTTTCTCATTTATGTTACT 2954  
Db 2901 TGAGGCTCTCTCTGGGTTTCTTAAACAAAGCTGGCTGATTTATCTCTCTGCTTCT 2960  
QY TCTAGTGTTCACCTCTGATTAAGATTTCTTGGTGAATAAGAAAAAATAAATAAATAAATA 3014  
Db 2961 ----GTGTACTTCTGTGATTAAGTCTCTTCTGGTGTCTAGAAAAAATAAATAAATA 3016  
QY 3015 AGGCGGCC 3022  
Db 3017 GGCAGGCC 3024

RESULT 9  
US-09-458-457-9  
; Sequence 9, Application US/09458457  
; Patent No. 6500654  
; GENERAL INFORMATION:  
; APPLICANT: Raju, Jeyaseelan  
; TITLE OF INVENTION: NOVEL CARP PROTEIN AND NUCLEIC ACID MOLECULES AND USES  
; FILE REFERENCE: mni-068CP  
; CURRENT FILING DATE: 1999-12-10  
; EARLIER APPLICATION NUMBER: 60/111,938  
; EARLIER FILING DATE: 1998-12-11  
; EARLIER APPLICATION NUMBER: 09/291,839  
; EARLIER FILING DATE: 1999-04-14  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 9  
; LENGTH: 2505  
; TYPE: DNA  
; ORGANISM: Rattus norvegicus  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(2505)  
US-09-458-457-9

Query Match 52.6%; Score 1892.2; DB 4; Length 2505;  
Best Local Similarity 84.9%; Pred. No. 0;  
Matches 212; Conservative 0; Mismatches 383; Indels 0; Gaps 0;  
QY 48 ATGGGAAATTAATAATCTAGACCAACCCAAATCTGTACTGTGAATGGAAGAAAAAGTC 107  
Db 1 ATGGGNAATTACAAATCCAGACCAACACAGACTTGTCTGTGAATGGAAGAAAGTT 60

QY 108 AGTGAATCATATGTTATCACAAATAGAAAGATTAGAAGATGACCTGCGAGATCAAGGAAAA 167  
DB 61 AGTGAATCTTACGCTATTATATAGAAAGGCTGAGGATAACCTGCGAGATCAAGGAAAAAT 120  
QY 168 GAACTGACAGAACTAAGAAATATATTTGGCTCTGATGAAGCCTTCAGTAAAGTCAATTTTA 227  
DB 121 GAAATTCAGAACTAAGGCAATCTTTGGCTCTGATGAAGCCTTCAGTAAAGTCAATTTTA 180  
QY 228 AATTACGCACTGAAAAATGSGCTGTCTCTACTTCAATTTATGTTGCAATTTGTGAGGCAAG 287  
DB 181 AATTACGCAAGAGCGTGGCTGTCCCTGTCTACACCTCTGCTGTCTGTGCGGCAAC 240  
QY 288 AATACACATATTCGAACCTTATGTTGAAAGGGCTCGCCCATCTCGACTGACAAAGAAAT 347  
DB 241 AAGTCAATATCCGTGCCCCATTATGTTAAAGGGCTCCGTCCATCCAGACTGACAGAAAT 300  
QY 348 GGATTTACAGCCTTGCAATTTAGCAGTTTACAGGATTAATGAGAAATGATCACTTCTCTG 407  
DB 301 GGGTTTCCAGCTCTGCACCTGGCCGTTTACAAAGACAGCCCGAACTTATCACTTCACTG 360  
QY 408 CTTACAGTGGAGCTGATATACAGCAGGTTGGATACGGTGGCCTCACTGCCCTCCATATT 467  
DB 361 TTGCAAGCGGAGCAGATGTTACAGCAAGTGGGATACGGTGGCCTCAAGCCCTCCACATA 420  
QY 468 GGTACAATAGCTGGCCACTAGAGGCTGCTGATGTCTGTTGCAACATGAGGATTAATGTC 527  
DB 421 GCTGCAATAGCTGGACACCCAGAGGCTGCAGAACTGCTGTACAAATGSGGCCAATGTG 480  
QY 528 AATATTCAGATGAGCTGTTTTTTTCACTCCATGATATTGAGCGTACTATGACATGAA 587  
DB 481 AATGTTCAAGATGCGCTCTCTTTCAACCCACTGCACATGACGCTACTATGGGACGAG 540  
QY 588 CAGGTAACCTCGCTCTCTTTTGAATTTGCTGCTGATGTAATGTAAGTGGTGAAGTTGGA 647  
DB 541 CAGGTAACAGCTGCTCTTTTGAATTTGCTGCTGATGTAATGTAAGTGGTGAAGTTGGA 600  
QY 648 GATAGACCCCTCCACTAGCATCTTGCAAAAGGATCTTGAATATTGCAAACTCTTGATG 707  
DB 601 GACAGGCTCTGCACCTGGCCTCTGCAAAAGGCTCTTCAACATTTGTGAACCTCTGGTA 660  
QY 708 GAAGAGGAGCAGACAGATGTAATGCTCAAGATTAATGAAGACCATGTCACACTCCAT 767  
DB 661 GAAGAGGAGCAGACAGATGTAATGCTCAAGATTAATGAAGACCATGTCACACTCCAT 720  
QY 768 TTCTGTTCTCGATTTGGAACCATGATATAGTTAAGTATCTGCTGCAAAAGTGAATGGA 827  
DB 721 TTCTGTTCTCGATTTGGAACCATGATATAGTTAAGTATCTGCTGCAAAAGTGAATGGA 780  
QY 828 GTTCAACCTCATGTTGTTAATATCTATGGAGATACCCCTTACACTGGCATGCTACAAT 887  
DB 781 GTCCAGCTCAGCTCATTAATCATATGTTGAGACACTCTTTGACACCTGGCATGCTACAAT 840  
QY 888 GGCATAATTTGAAGTTGCCAAGAAATCATCCAAATATCAGAAATCAGAAAGTCTGACTAAG 947  
DB 841 GGAATTTTGAAGTTGCCAAGAAATTTGTCAGGTAACAGAACTGAAAGTCTGACTAAG 900  
QY 948 GAAACATCTTCAGTGAACAGCTTTTCAAGTCTGTTGACCTATGCGAAGCATGATGAC 1007  
DB 901 GAAACATCTTCAGGAGACAGCTTTTCAAGTCTGTTGACCTATGCGAAGCAATGAC 960  
QY 1008 CTAGTCAAAATTTCTTTGATCAGAATGTATAAACATCAACACCAAGGAGGATGGG 1067  
DB 961 CTGGTCAAAATTTCTTTGATCAGATGCTGTAACATTAACACCGAGGAGAGATGGG 1020  
QY 1068 CACATGGAATFACATCTGCTGTGTAACACCGGTCAATTCGCTGGTTCAAGTCTTACTG 1127  
DB 1021 CACACAGGATTCACATCTGCTGTGTACCAAGCCCATATCCGCTGGTTCAAGTCTTACTT 1080  
QY 1128 GATAATGAGCTGATATCAATCTAGTGGCTTGTGATCCAGCAGGCTCTAGTGGTGA 1187  
DB 1081 GATAATGAGCTGATATCAATCTAGTGGCTTGTGATCCAGCAGGCTCTAGTGGTGA 1140  
QY 1188 GATGAGCAGACATGTTTGTGATGCGGCTTTATGAAAAAGGCGCATGATGCCATTTGACACTC 1247

DB 1141 GATGAGCAGACATGTTTGTGATGTTGGCTTACGAGAAAGGACATGATGCCATTGTTACACTC 1200  
QY 1248 CTGAAGCATTTATAGAGACCAAGAATGAAATGCTCCCTGTAAATGAAATTTCTCAGGCTGGA 1307  
DB 1201 CTGAAGCATTTACAAAGACACCCAGAGGAGCTGCCATGTAAGAAATTTCCAGGCTGGA 1260  
QY 1308 GGAGATGGCTCCTATGTTGTTGTTCCATCACCTTGGGGAAGATTAAAAAGCATGACAAA 1367  
DB 1361 GGAGATGGCTCCTATGTTGTTGTTCCCTTCCCTTGGGCAAGATTAAAAAGCATGACAAA 1320  
QY 1368 GAGAAGCGAGATATTTCTCTCTAAGAGCTGGAATGCTTCAATTTCAATTTCAAGTCTC 1427  
DB 1321 GAGAAGCGAGATGTTCTCTCTCTGAGGCTGAATACCTCCGCTTCCATCTCCAACTC 1380  
QY 1428 TCAGAAATTCAGTTCCATGAGATTATGGCTCAGGTTCTTTTGGGAAAGTATATAAGGA 1487  
DB 1381 TCCGAAATTCAGTTCCACGAGATTATCGGCTCGGTTCTTTTGGGAAAGTCTATAAGGG 1440  
QY 1488 CGATGCAGAAATAAAAATAGTGGCTATAAAACGTTATCGAGCCAAATACCTTACTGCTCCAAG 1547  
DB 1441 CGATGCAGAAATAAAAATAGTGGCGATCAAAACGATACCGAGCCAAACCTTACTGCTCCAAG 1500  
QY 1548 TCAGATGTGATATGTTTTGCCGAGAGTGTCCATTTCTGCGAGCTCAATCATCCCTGC 1607  
DB 1501 TCAGACGTGATATGTTTTGCCGAGAGTGTCCATTTCTGCGAGCTCAACACCCCTGC 1560  
QY 1608 GTAATTCAGTTTGGGTGCTTCTTGAATGATPCCAGCAGTTTGGCATTGCTCACTCAA 1667  
DB 1561 GTGGTTCAAGTTTGGGTGCTGCTGATGATGACCCAGTCAGTTTGGCATTGCTCACTCAG 1620  
QY 1668 TACATATCAGGGGTTCTGTTCTCCCTCTCTCATGAGCAGAGAGATTCTTGAATTG 1727  
DB 1621 TACATTTCAAGGGCTCCCTGTTCTCCCTGCTTCATGAACAGAGAGATTCTTGAATTG 1680  
QY 1728 CAGTCTAAATTAATTTGAGTAGATGTTGCCAAAGGATGGAGTACCTTCAACACCTG 1787  
DB 1681 CAGTCTAAATTAATTTGAGTAGATGTTGCCAAAGGATGGAGTACCTTCAACACCTG 1740  
QY 1788 ACACAGCAATATATACATCGTACTTTGAACAGTCACAATATTTCTTCTATGAGATGGG 1847  
DB 1741 ACCCAGCAATATATACACCCGACCTGAAACAGCCAAATATTTCTGCTCTATGAGATGGC 1800  
QY 1848 CATCTGTGTGGCAGATTTTGGAGATCAAGATTTCTACAGTCTCTGATGAAGACACAC 1907  
DB 1801 CATCTGTGTGGCAGATTTTGGAGATCAAGATTTCTGAGTCCCTGGATGAAGACACAC 1860  
QY 1908 ATGACAAAACAACTCGGAACTCCGTTGGATGGCTCCTCAGGTGTTTCAAGCAGTGCCT 1967  
DB 1861 ATGACAAAACAACTCGGAACTCCGTTGGATGGCTCCTCAGGTGTTTCAACAGTGCAGG 1920  
QY 1968 CGGTACACCATCAAGCAGATGTTTCAAGTATGCTCTGTGTCTGTGGGAAATTTCTCACT 2027  
DB 1921 AGATACACCATCAAGGCTGATGTTCTTCAAGTATGCTCTGTGTCTGTGGGAGTCTCTCACT 1980  
QY 2028 GGCAGAAATTCATTTCCGCTCATCTCAAGCCAGCGGCTCGCGCAGACAGATGCTTACCAAC 2087  
DB 1981 GGAGAAATTCATTTCCGCTCATCTCAAGCCAGCGGCTCGCGCAGACAGATGCTTACCAAC 2040  
QY 2088 CACATCAGACTCCCATTTGGCTATTCATTTCCCAAGCCCATATCATCTCTGTGTATGGA 2147  
DB 2041 CACATCAGCGGCTTCCGCTATTCATTTCCCAAGCCCATCTCATCTCCCTGTGTATGCGG 2100  
QY 2148 GGGTGAACCATGCTCTGAAGGAGAACCCGAAATTTTCTGAAGTGTCTCATGAAGTTAGAA 2207  
DB 2101 GGCTGGAATGCAATGCTCTGAAGGAGAACCCGAAATTTTCTGAAGTGTCTCATGAAGTTAG 2160  
QY 2208 GAGTGTCTCTGCAACATGAGCTGATGTTCTCTGCACTCAAGTAAACAGCAGTGGGTCTCTC 2267  
DB 2161 GAGTGTCTCTGCAACATGAGCTGATGTTCTCTGCACTCAAGTAAACAGCAGTGGGTCTCTC 2220  
QY 2268 TCACCTTCTTCTTCTGATTCGCTGGTGAACCCGGGAGGACCTTGGCGGAGTCTATG 2327

Db 2221 TCACCTTCTCTTCTCCGATTGCTGCTAGCGGGGGGCGCTGGCGGAGCCACGTG 2280  
Qy 2328 GCAGCATTAAGAAAGTCTTTTCGAATTTGGAATATGCTCTAAATGCAAGGTCTCTATGCTGCT 2387  
Db 2281 GCAGCTTACGGAGCGCTTTTGAGTTGGAGTATGCCCTAAATGCAAGGTCTCTATGCTGGG 2340  
Qy 2388 TTGTCCTCAAGTGTGTCGACAAATTCCTCTCAAGGTCTCTCTTTGGAGGAGATGAAAAGA 2447  
Db 2341 TGGTCCCAAGTGTGTCGACAACTCTAATCCGGGCTGCTCTTTGGAGGAGATGAATAGG 2400  
Qy 2448 AGTCTTCAATACACACCATTTGACAAATATGGCTATGATCCGATCCCATGAGCTCAATG 2507  
Db 2401 AGACCCAGTATTCAATGTTGACAAATACGGCTATGTTCTGATCCCATGAGCCTGAGC 2460  
Qy 2508 CATTTTCATTTCTGCCAAATAGTAGCAGCTTTGAGACAGCAGC 2552  
Db 2461 CACCTTCACTCCCGCAAGACGACGACAACTTTTGGAGACAGAAC 2505

RESULT 10  
US-09-947-199A-9  
; Sequence 9, Application US/09947199A  
; Patent No. 6660490  
; GENERAL INFORMATION:  
; APPLICANT: Raju, Jeyaseelan  
; TITLE OF INVENTION: NOVEL DARK PROTEIN AND NUCLEIC ACID MOLECULES AND USES  
; FILE OF INVENTION: THEREFOR  
; FILE REFERENCE: MNI-068CP2  
; CURRENT APPLICATION NUMBER: US/09/947,199A  
; CURRENT FILING DATE: 2001-09-05  
; PRIOR APPLICATION NUMBER: US 60/111,938  
; PRIOR FILING DATE: 1998-12-11  
; PRIOR APPLICATION NUMBER: US 09/291,839  
; PRIOR FILING DATE: 1999-04-14  
; PRIOR APPLICATION NUMBER: US 09/458,457  
; PRIOR FILING DATE: 1999-12-10  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 9  
; LENGTH: 2505  
; TYPE: DNA  
; ORGANISM: Rattus norvegicus  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)...(2505)  
US-09-947-199A-9

Query Match 62.6%; Score 1892.2; DB 4; Length 2505;  
Best Local Similarity 84.7%; Pred. No. 0;  
Matches 2122; Conservative 0; Mismatches 383; Indels 0; Gaps 0;

Qy 48 ATGGGAATTAATCTAGACCAACCCAACTTGCTGATGATGGAAGAAAAGTC 107  
Db 1 ATGGGGAATTAACAAATCCAGCAACACAGACTTGTCTGATGAATGGAAGAAAAGTT 60  
Qy 108 AGTGAATCATATGTTATCAACATAGAAAAGATTAGAAGATGACCTGCAGATCAAGAAAATA 167  
Db 61 AGTGAATCTTACGCTATTATCATAGAAAGGCTGGAGGATAACCTGCAGATCAAGAAAAT 120  
Qy 168 GAATGACAGAACTAAGGAATATATTGGCTCTGATGAAGCCCTTGATGAAGTCAATTTA 227  
Db 121 GAATTTCAAGAACTAAGGCACATCTTTGGCTCTGATGAAGCCCTTCAAGTCAAGTTA 180  
Qy 228 AATTACCGCACTGAAATGGGCTGCTCTACTTCTTATGTTGATTTGTCGAGGCAAG 287  
Db 181 AATTACCGCAGACAGCGTGGCCCTGCTCCTGCTACACTCTGCTGCTGCTGCTGGCGGCAAC 240  
Qy 288 AAATCACATATCGAACTCTTATGTTGAAAGGCTCGGCCCATCTCGACTGACAGAAAT 347  
Db 241 AAGTCATATCCGTCGCCCTTATGTTAAAGGGCTCCGTCCTCCATCCAGACTGACGAGAAAT 300  
Qy 348 GGATTTACGCTTGCATTTAGCAGTTTACAAGGATTAATCCAGATGATGATCACTTCTCTG 407

Db 301 GGGTTTTCCAGCTCTGCACCTGGCCGTTTTACAAAGACAGCCCGAACTTTATCACTTCACTG 360  
Qy 408 CTTTCACAGTGGAGCTCATATACAGCAGGTTGGATACGGTGGGCTCACTGCCCTCCATATT 467  
Db 361 TTGCAACAGCGGACGATGTTGACAAAGTGGGATACGGTGGGCTCACAGCCCTCCACATA 420  
Qy 468 GTTCAATAGCTGGCCACCTAGAGGCTGCTGATGCTGCTTTGCCAACAATGAGAGCTAATGTC 527  
Db 421 GCTGCAATAGCTGGACACCCAGAGGCTGAGAAAGTGTCTTACAACTGGGGCCAAATGTG 480  
Qy 528 AATATTCAAGATGCACTTTTTTTTCACTCCATTCATATTTGACAGCTTACTATGACATGAA 587  
Db 481 AATGTTCAAGATGCGCTCTTTTCCACCACCTGCACATTTGAGCTTACTATGGGACGAG 540  
Qy 588 CAGGTAACCTGCGCTTTTGTGAAATTTGCTGCTGATGTAATGTAAGTGGTGAAGTTGGA 647  
Db 541 CAGGTAACCAAGTGTCTTTTGAAGTTTGGTGTGATGTCAATGTAAAGCGGTGAAGTTGG 600  
Qy 648 GATAGACCCCTCCACCTAGCATCTGCAAAAGGATCTTGAATATTGCAAAACTCTTGATG 707  
Db 601 GACAGGCTCTGCACCTGGCCCTTGCAAAAGGGCTTCTTCAACATTTGAAACTCTCTGTA 660  
Qy 708 GAAAGAGGACGACAAAGCAGATGTGAATGCTCAAGATAATGAAGACCATGTCCCACTCCAT 767  
Db 661 GAAGAGGGAGCAAAACAGATGTGAACGCTCAGGACATGAAGACCACTCCCTCTGAC 720  
Qy 768 TTTGTTCTCGATTTGACACCAATGATAGTATAGTATCTGCTSCAAAGTGAATTTGAA 827  
Db 721 TTTGTTCTCGATTTGACACCAATATAGTAGTACCTGCTCCAGAGTGAATTTAGAG 780  
Qy 828 GTTCAACCTCATGTTGTAATATCTATGAGATACCCCTTACAGCTGACATGGCATGTACAAT 887  
Db 781 GTCCAGCTCACGCTAATTAACATCTATGATGATCTCTTTGCACTCTGGCATGCTACAAT 840  
Qy 888 GGCATAATTTGAAGTTGCCAAGGAAATCATCCAAATATCAGAAACAGAAAGTCTGACTAAG 947  
Db 841 GGAATTTTGAAGTTGCCAAGGAAATTTGCCAGGTAACAGGAACTGAAAGTCTGACTAAG 900  
Qy 948 GAAACATCTTTCAGTGAACACAGCTTTTCATAGTGTGTGACCTATGCGAAGAGACTGAC 1007  
Db 901 GAAACATCTTTCAGGAGACAGCTTTTCACAGTGTGTGACCTATGCGAAGAACTGAC 960  
Qy 1008 CTAGTCAAAATTTCTTCTGATCAGAAATGTATAAATCAACCAACCAAGGAGGATGGG 1067  
Db 961 CTGTCAAAATTTCTTCTGATCAGAAATGTGTGAACATTAACCAACCGAGGAAAGATGG 1020  
Qy 1068 CACACTGGATTACACTCTGCTTACAGGCTGACATTCGCGCTGGTTCAGTTCTTACTG 1127  
Db 1021 CACACAGGATTGCACTCTGCTTGTCTACACGCGCATATCCGCTGGTTCAGTTCTTACTT 1080  
Qy 1128 GATAATGGAGCTCATATGAATCTAGTGGCTTGTGATCCCGAGAGCTTAGTGGTGAATAA 1187  
Db 1081 GATAATGGTGCAGATATGAATCTTGTGCTTGTGATCCCGAGAGCTTAGTGGTGAATAA 1140  
Qy 1188 GATGAGCAGACATGTTTGTGATGTTGGCTTATGAAAAGGCGATGATGCCATTTGACATC 1247  
Db 1141 GATGAGCAGACATGTTTGTGATGTTGGCTTACGAAAAGGACATGATGCCATTTGACATC 1200  
Qy 1248 CTGAGCATTTAAGAGACCAACAGATGAATTCCTGTAATGAATATTCTCAGGCTGGA 1307  
Db 1201 CTGAGCATTTAAGAGACCAACAGGAGCTGCCATGTAACGAATATTTCCAGGCTGGA 1260  
Qy 1308 GGAGATGGCTCTCTATGCTGTTTCCATCACCTTTGGGAAAGATTAAAGCATGACAAA 1367  
Db 1261 GGAGATGGCTCTCTATGCTGTTCTTCCCTTTGGGCAAGATTAAAGCATGACAAA 1320  
Qy 1368 GAGAGGCGAGATTTCTCCTTCAAGAGCTGATTTGCTTCACTTCCATTTCCATTTCCAGTC 1427  
Db 1321 GAGAGGCGAGATTTCTCCTTCAAGGCTGAACTACCTCCCGCTTCCATTTCCATTTCCACTC 1380  
Qy 1428 TCAGAAATTTGAGTTCCATGATTAATTTGGCTCAGGTTCTTTTGGGAAAGTATATAAGGA 1487  
Db 1381 TCCGAATTCAGTTCCACGAGATTTACGGCTCGGTTCTTTGGGAAAGTCTATATAAGGG 1440

```
QY 1488 CGATCAGAAATAAATAGTGGCTATAAACGTTATCGAGCCAAATCACTACTGCTCCAAG 1547
Db 1441 CGATCAGAAATAAATAGTGGCGATCAACGATACCGAGCCAACTACTACTGCTCCAAG 1500
QY 1548 TCAGATGTGGATATGTTTGGCGAGAGGTGTCATCTCTGCCAGCTCAATCATCCTGTC 1607
Db 1501 TCAGACGTGGATATGTTTGGCGAGAGGTGTCATCTCTGCCAGCTCAACACCCTGTC 1560
QY 1608 GTAATTCAGTTTGTGGGTGCTTGTGTAATGATCCAGCCAGTTTGGCCATTTGTCACAA 1667
Db 1561 GTGGTTCAAGTTTGTGGGTGCTTGTGTAATGATCCAGCCAGTCAATTTGTCACCTCAG 1620
QY 1668 TACATATCAGGGGTTCTGTTCTCCCTCTTCATGAGCAGAGAGGATTCCTGATTG 1727
Db 1621 TACATTTCAGGAGGCTCCCTGTTCTCCCTGCTTCATGAACAGAGAGAATTCCTGACTG 1680
QY 1728 CAGTCTAAATTAATTCGAGTAGATGTTGCCAAAGCATGGAGTACCTTCACAACTG 1787
Db 1681 CAGTCTAAATTAATTCGAGTAGAGTTCGCCAAGGCATGGAGTACCTTCGACAGTTG 1740
QY 1789 ACACAGCAATTAATCATCGTGACTTGAACAGTCAATATTTCTTCTATGAGGATGGG 1847
Db 1741 ACCCAGCAATCATACACCGGACCTGAACAGCCACAATATTTCTGCTCTATGAGGATGGC 1800
QY 1848 CATGCTGTGGTGGCAGATTTTGGAGAAATCAAGATTTCTACAGTCTCTGGATGAAGCAAC 1907
Db 1801 CATGCTGTGGTGGCAGATTTTGGAGAAATCAAGATTTCTGAGTCCCTGGATGAAGCAAC 1860
QY 1908 ATGACAAACAACTGGGAACTCCGTTGGATGCTCTGAGGTGTTTCACGAGTGCACT 1967
Db 1861 ATGACAAACGACCCAGGAACTCCGTTGGATGCTCTGAGGTGTTTCACAGTGCACT 1920
QY 1968 CGGTACACCATCAAGCAGATGCTTCAGCTATGCTCTGTTGCTGTGGGAAATTCCTCACT 2027
Db 1921 AGATACCATCAAGCCTGATGCTTCAGTTACTTCCCTGTTCTGTGGAGTCTCTCACT 1980
QY 2028 GCGGAAATTCATTCGCTCATCTCAAGCCAGCGGCTCGGCGAGCAGACATGGCTTACCAC 2087
Db 1981 GGAGAAATTCATTCGCTCATCTCAAGCCAGCGGCTCGGCGAGCAGATATGGCGTATCAC 2040
QY 2088 CACATCAGACCTCCCATTTGGTATTTCCATTCGAGGCCATATCATCTCTGCTGATACGA 2147
Db 2041 CACATCAGACCGCCCATCGGCTATTCATCCGAGCCCATCTCATCCCTGCTGATACGG 2100
QY 2148 GGGTGGAAAGCATGCTCGTGAAGGAGACCCGAAATTTCTGAAGTTGTCATGAAGTTAGAA 2207
Db 2101 GGTGGAATGCATGCTGAAGGACGACAGAGTTCTGAGTCTGTAGCNAACTGGAG 2160
QY 2208 GAGTGTCTGCAACATTTGAGTCTCTGATCAATCAAGTAAACAGCAGTGGGTCTCTC 2267
Db 2161 GAGTGCCTATGCAATGTGGAGTCAATGCTCTGAGCAGTCAAGTAAACAGCAGTGGGTCTCTG 2220
QY 2268 TCACCTCTCTCTCTCTCTGATTTGCTGTAACCGGAGAGACCTGCGGAGTCTATGTCG 2327
Db 2221 TCACCTCTCTCTCTCTCTGATTTGCTGTAACCGGAGAGGCTGCGGAGAGTCTATGTCG 2280
QY 2328 GCAGCATTAAGAAGTCGTTTCCAAATGGAATATGCTCTAAATGAAGTCTTATGCTGT 2387
Db 2281 GCAGCCTTACGAGGCGGTTTGGTGGAGTATGCCCTAAATGAAGTCTTATGCTGTGGG 2340
QY 2388 TTGTCGCAAGTCTGGACATATTCCTCTCAAGTCTGCTTTGGAGGAGATGAAGA 2447
Db 2341 TGGTCCCAAGGTTTGGAAACACATCTAAATCCGGGCTGCTTTGGAGGAGATGAATAGG 2400
QY 2448 AGTCTTCAATACACACCCATTTGACAAATATGGCTATGATCCGATCCATGAGCTCAATG 2507
Db 2401 AGCACCAGATTTCAACTGTTGACAAATACGGCTATGCTGTGATCCATGAGCCTGAGC 2460
QY 2508 CATTTTCAATCTTGGCGGAAATAGTAGCAGCTTTGAGGACAGCAGC 2552
Db 2461 CACCTTCACTCCCGCCAAAGACGACAGCAACTTTTGGAGCAGCAAC 2505
```

```
RESULT 11
US-09-833-381-1151
; Sequence 1151, Application US/09833381
; Patent No. 6672186
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; PRIORITY FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1151
; LENGTH: 740
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(740)
; OTHER INFORMATION: n = A,T,C or G
US-09-833-381-1151

Query Match 14.1%; Score 427.6; DB 4; Length 740;
Best Local Similarity 78.4%; Pred. No. 7.4e-107;
Matches 519; Conservative 0; Mismatches 142; Indels 1; Gaps 1;

QY 16 CGGCCCTGGAGAAAGGAGAACTTATATAAATGGGAAATTTATAATCTACCAACCC 75
Db 41 CAGCACAGGAGAAAGCAAGACTTCTTTAAATGGGAAATTAACAATCCAGCAACAC 100
QY 76 AAACCTGTCTGATGAATGAAGAAAGTCAAGTGAATCATATGTTATCAAAATAGAAA 135
Db 101 AGACTTGTCTGATGAATGAAGAAAGTGTAGTGAATCTTACGCTATTAATCATAGAAA 160
QY 136 GATTAGAGATGACCTGCAGATCAAGGAAAGAAAGTGAAGTGAAGTGAAGTGAAGTGA 195
Db 161 GGCTGGAGGATAACCTGCAGATCAAAAGAAATGAATTTCAAGAACTTAAGGCAATCTTTG 220
QY 196 GCTCTGATGAAGCTTCAGTAAAGTCAATTTAAATTAACCGCACTGAATATGGCTGTCTC 255
Db 221 GCTCTGATGAAGCTTCAGTAAAGTCAATTTAAATTAACCGCACTGAAGTGAAGTGAAGT 280
QY 256 TACTTCAATTTATGTTGCAATTTGTGGAGGCAAGAAATCAATATTCGAACTTTATGTTGA 315
Db 281 TGCTACACCTCTGCTGTGCTGTGGCGCAACAAGTCAATATCGTGCCCTTATGTTAA 340
QY 316 AAGGGCTCCGCCATCTCGACTCAAGAAATGGATTTACAGCTTGCATTTAGCAGTTT 375
Db 341 AAGGGCTCCGCCATCTCGACTCAAGAAATGGATTTACAGCTTGCATTTAGCAGTTT 400
QY 376 ACAAGGATAATGAGAAATTCATCTTCTGCTTCACT-AGTGGAGCTGATATACAGCAG 434
Db 401 ACAAGGACAGCCGGAATTCATCTTCACTGTTGCAAGCGGAGCAGATGTTTCAGCAA 460
QY 435 GTTGGATACGGTGGCTCTACCTGCCCTCCATATTTGCTAATAGTGGCCACCTAGAGCT 494
Db 461 GTGGGATACGGTGGCTCTACAGCCCTCCACATAGCTGCAATAGCTGGACACCCAGAGCT 520
QY 495 GCTGATGCTGTGTTGCAACATGGAGTAAATGTTCAATATTAAGATGCAAGTGTCTTCTACT 554
Db 521 GCANAAGTGTGNTGCAACATGGGCGCAACGTCATGTTCAAGATGCGCTCTTCTTACC 580
QY 555 CCATTTGATATTCAGCGTACTATGAGCATGAAGCAGGTAACCTGGCTCTTTTGAATTT 614
Db 581 CCATGTCACATTCAGCGCTACTATGGGCGCAGCAGGTAACCGAGTGTCTCTTTGAAGTT 640
QY 615 GGTGCTCATGTAATGTAAGTGTGAAGTTGGAGATAGACCCCTCCACCTAGCATCTGCA 674
Db 641 GGNGGTGATGTCNATGTAAGCGGTNAAGTTGGNACAGGCTCCGCCCTCGCCNCGNA 700
```

QY 675 AA 676  
||  
Db 701 AA 702

## RESULT 12

US-09-833-381-1149  
; Sequence 1149, Application US/09833381

; Patent No. 6672186

; GENERAL INFORMATION:

; APPLICANT: Robison, Keith E.

; TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs

; FILE REFERENCE: 5800-119

; CURRENT APPLICATION NUMBER: US/09/833,381

; CURRENT FILING DATE: 2001-04-11

; PRIOR APPLICATION NUMBER: 09/516,448

; PRIOR FILING DATE: 2000-02-29

; NUMBER OF SEQ ID NOS: 2050

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 1149

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-833-381-1149

Query Match 14.1%; Score 427.4; DB 4; Length 616;

Best Local Similarity 81.8%; Pred. No. 7.6e-107;

Matches 505; Conservative 0; Mismatches 111; Indels 1; Gaps 1;

QY 338 GACAAGAAATGGATTTACAGCTTGTGATTTAGCAGTTTCAAGAGTAATGCGAATTGAT 397  
|||  
Db 1 GACAGAAATGGGTTTCCAGCTCTGCACCTGGCGGTTTCAAGGACAGCCCGGAATTAT 60

QY 398 CACTTCTCTGCTTACAGTGGAGCTGATATACAGCAGGTTGGATACGGTGGCTCACTGC 457  
|||  
Db 61 CACTTCACTGTGACAGCGGAGCAGATGTTTCAGCAAGTGGGATACGGTGGCTCACAGC 120

QY 458 CCTCCATATGCTACATAGCTGGCCACCTAGAGGCTGTGATGCTGCTTCCACATGG 517  
|||  
Db 121 CCTCC-CATAGCTGCAATAGCTGACACCCAGAGGCTGCAAGAGTCTCTGCAACATGG 179

QY 518 AGCTAAATGTCAAATATTCAGATGAGTGTCTTTTCACTTCCATTTGCAATATTCGACGCTACTA 577  
|||  
Db 180 GGCCACAGTGAATGTTCAAGATGCGGTCTTCTTCCACCCACTGCACATTCGACGCTACTA 239

QY 578 TGGACATGAACAGATGACTCGCTTCTTTTGAATTTGGTGTGATGTAATGTAAGTG 637  
|||  
Db 240 TGGGCACGACGAGTAAACCACTGTCTTTTGAAGTTTGGTGTGATGTCATGTAAGCGG 299

QY 638 TGAAGTTGGAGATAGACCCCTCCACCTAGCATCTGCAAAAGGATTTCTTGAATATTCGAA 697  
|||  
Db 300 TGAAGTTGGGACAGGCTCTGCACCTGGGCTTTCGAAGGGGTTCTTCAACATTTGAA 359

QY 698 ACTCTTGATGAAGAAGCGCAGCAAGCAGATGTAATGCTTCAAGATPAATGAAGACCATGT 757  
|||  
Db 360 ACTCCTGTTGAAGAAGGAGCAAGCAGATGTAACGCTTCAGCAATGAAGACCCAGCT 419

QY 758 CCCACTCCATTTCTGTTCTGATTTGGACACCATGATATAGTTAACTATCTGCTGAAG 817  
|||  
Db 420 CCCCTGACCTTCTGTTCTGATTTGGACACCAATATAGTAGGCTACTGCTCCAGAG 479

QY 818 TGATTTGGAAGTTCAACCTCATGTTGTTAATATCTATGGAGATPACCCCTTACACCTGCG 877  
|||  
Db 480 TGACTTAGAGGTCCAGGCTCAGGTCATTAAACATCTATGGTGAACACTCTTTGCACTGCG 539

QY 878 ATGCTACAAATGGCAAAATTTGAAGTTGCCAAGGAATCATCCAATATCAGGAACAGAAG 937  
|||  
Db 540 ATGCTACAAATGGAAATTTGAAGTTGCCAAGGAATTTGTCAGGTAACAGGAACCTGAAG 999

QY 938 TCTGACTAAGGAAACA 954  
|||

Db 500 TCTGACTAAGGAARACA 616

## RESULT 14

US-09-833-381-1147/c

; Sequence 1147, Application US/09833381

## RESULT 13

US-09-833-381-1150  
; Sequence 1150, Application US/09833381

; Patent No. 6672186

; GENERAL INFORMATION:

; APPLICANT: Robison, Keith E.

; TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs

; FILE REFERENCE: 5800-119

; CURRENT APPLICATION NUMBER: US/09/833,381

; CURRENT FILING DATE: 2001-04-11

; PRIOR APPLICATION NUMBER: 09/516,448

; PRIOR FILING DATE: 2000-02-29

; NUMBER OF SEQ ID NOS: 2050

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 1150

; LENGTH: 736

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc\_feature

; LOCATION: (1)..(736)

; OTHER INFORMATION: n = A,T,C or G

US-09-833-381-1150

Query Match 13.7%; Score 414.6; DB 4; Length 736;

Best Local Similarity 81.2%; Pred. No. 2.6e-103;

Matches 480; Conservative 0; Mismatches 111; Indels 0; Gaps 0;

QY 312 TTGAAGGGCTCCGCCCATCTCGACTGACAAGAAATGGATTTACAGCCCTTGCATTTAGCA 371  
|||  
Db 1 TTAAAGGGCTCCGTCATCCAGACTGACGAGAAATGGGTTCCAGCTCTGCACCTGGCC 60

QY 372 GTTTACAAGGATTAATGAGAAATGATCACTTCTCTGCTTCAAGTGGAGCTGATATACAG 431  
|||  
Db 61 GTTTACAAGGACAGCCGGAACTTATCACTTCACTGTTGCACAGCGGAGCAGATGTTCA 120

QY 432 CAGTTTGGATACGGTGGCTCTCAGTCCCTCCATTTGCTTACATAGCTGSCACCTAGAG 491  
|||  
Db 121 CAGTTGGGATACGGTGGCTCTCAGTCCCTCCATTTGCTTACATAGCTGSCACCTAGAG 180

QY 492 GCTGCTGATGTGCTTGTGCAACATGGAGCTTAATGTCAATATTTCAAGATCAGTTTTTTC 551  
|||  
Db 181 GCTGCAAGAGTGTCTGTACAAATGGGCAATGTGAATGTTCAAGATCCCGTCTTCTTC 240

QY 552 ACTCCATTCATATTCGAGCGTACTATGACATGAACAGTAACTGCGCTTCTTTGAA 611  
|||  
Db 241 ACCCACTGCACATTCGAGCTACTATGGGCAAGCAGGTAAACCAGTGTCTCTTTGAAG 300

QY 612 TTTGCTGCTCATGTAAATGTAAGTGTGAAGTTGGAGATAGACCCCTCCACCTAGCATCT 671  
|||  
Db 301 TTTGCTGCTCATGTCAATGTAAAGTGTGAAGTTGGGACAGGCTCTGACCTGGCCTCT 360

QY 672 GCAAAAGGATTTCTGATATTTGCAAACTCTTTGATGGAGAGGCGCAAGCAGATGTG 731  
|||  
Db 361 GCAAAAGGCTTCTTCAACATTTGTGAACCTCTGTTAGAAAGGAGGAGCAAGCAGATGTG 420

QY 732 AATGCTCAAGATTAATGAAGCCCATGTCCTCACTCCATTTCTGTTCTCGATTTGGACACCAT 791  
|||  
Db 421 AACCTCAGGACCAATGAAGACCACTGTCCTCTGCACTTCTGTTCTCGATTTGGACACCAT 480

QY 792 GATATAGTTAAAGTATCTGCTGCAAGTGAATTTGGAAGTTCAACCTCATGTTGTTAATATC 851  
|||  
Db 481 AATATAGTAGCTACCTGCTCCAGAGTGACTTANAGGTTCCAGCTCCAGCTCATTAACATC 540

QY 852 TATGGAGATACCCCTTACACCTGGCATGCTACAAATGGCAATTTGAAGTT 902  
|||

Db 541 TATGGTGACACTCCTTTGACCTGGCATGCTCAATGGAATTTTGAANTT 591  
|||

Patent No. 6672186  
GENERAL INFORMATION:  
APPLICANT: Robison, Keith E.  
TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs  
FILE REFERENCE: 5800-119  
CURRENT APPLICATION NUMBER: US/09/833,381  
CURRENT FILING DATE: 2001-04-11  
PRIOR APPLICATION NUMBER: 09/516,448  
PRIOR FILING DATE: 2000-02-29  
NUMBER OF SEQ ID NOS: 2050  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 1147  
LENGTH: 304  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc.feature  
LOCATION: (1)...(304)  
OTHER INFORMATION: n = A,T,C or G  
US-09-833-381-1147

Query Match 6.1%; Score 184.2; DB 4; Length 304;  
Best Local Similarity 79.6%; Pred. No. 1.4e-40;  
Matches 242; Conservative 0; Mismatches 59; Indels 3; Gaps 2;

QY 313 TGAAGGGCTCGCCCATCTCGACTGACAGAAATGGATTTACAGCC--TTGCATTAGC 370  
DB |||||  
304 TAAAGGGCTCGCTCATCCAGACTGACGAGAAATGGTTTCCAGATTTTGACCTGCC 245  
QY 371 AGTTTACAGGATAATGAGAAATGATCAC-TTCTGTCTTCAAGTGGAGCTGATATAC 429  
DB |||||  
244 GGTTTACAGGACAGCCCGGAATTTATCACTTTCACTGTTCACAGCGGAGCAGATGTT 185  
QY 430 ACAGAGTGGATACGGTGGCTCACTGCCCTCCATATGCTACATAGCTGGCCACTAG 489  
DB |||||  
184 AGCAAGTGGATACGGTGGCTCACTGCCCTCCATAGCTGCAATAGCTGGACACCCAG 125  
QY 490 AGGCTGCTGATGTGTTGTCACAAATGGAGCTAATGTCAATATTCAGATGCAAGTTTTT 549  
DB |||||  
124 AGGCTGCAAGAGTGTGATGCAACATGGGCGCAACGTGAATGTTCAAGATGCCGTCTCT 65  
QY 550 TCACTCCATTGCATATTCAGGCTACTATGACATGACAGGTAAGTTCGCTTCTTTTGA 609  
DB |||||  
64 TCACCCCACTGCATTCAGCTACTATGGGCGACGAGGTAACCAAGTGTCTCTTTTGA 5  
QY 610 AATT 613  
DB 4 AGTT 1

RESULT 15  
US-09-833-381-1148/c  
Sequence 1148, Application US/09833381  
Patent No. 6672186  
GENERAL INFORMATION:  
APPLICANT: Robison, Keith E.  
TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs  
FILE REFERENCE: 5800-119  
CURRENT APPLICATION NUMBER: US/09/833,381  
CURRENT FILING DATE: 2001-04-11  
PRIOR APPLICATION NUMBER: 09/516,448  
PRIOR FILING DATE: 2000-02-29  
NUMBER OF SEQ ID NOS: 2050  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 1148  
LENGTH: 186  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc.feature  
LOCATION: (1)...(186)  
OTHER INFORMATION: n = A,T,C or G  
US-09-833-381-1148

Query Match 2.8%; Score 84.4; DB 4; Length 186;  
Best Local Similarity 69.1%; Pred. No. 1.9e-13;  
Matches 112; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 452 CACTGCCCTCCATATTCCTACAATAGCTGGCCACCTAGAGGCTGCTGTGATGCTGTGCA 511  
DB |||||  
162 CCCTTCCAANATAGTCCCAATAGTTGNACACCCCAANAGGTGCAAAAGTCTGTGTC 103  
QY 512 ACATGGAGCTAATGCTCAATATTCAGATGCAAGTTTTTTTCACTCCATTCATATTCAGC 571  
DB |||||  
102 ACATGGGCGCAATGTGAATGTTCAAGATGCCGCTCTTTCACCCCACTGCACATTCAGC 43  
QY 572 GTACTATGGACATGAACAGGTAACCTGCCCTTCTTTTGAATT 613  
DB |||||  
42 CTACTATGGGCGAGGAGGTAACCAAGTGTCTCTTTTGAAGTT 1

Search completed: September 6, 2004, 07:42:49  
Job time : 152.506 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 6, 2004, 05:03:46 ; Search time 915.158 Seconds  
(without alignments)  
16430.147 Million cell updates/sec

Title: US-10-626-173-1

Perfect score: 3025

Sequence: 1 GTGACCCAGCGTCGGCCCTGGAGAAAGAACTTATATAATGGAATTATA 60

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 3267054 seqs, 2485319735 residues

Total number of hits satisfying chosen parameters: 6534108

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:

1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq.\*  
12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*  
13: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*  
14: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*  
16: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq.\*  
17: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*  
18: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*  
19: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3025	100.0	3025	9	US-09-947-199-1
2	3025	100.0	3025	17	US-10-626-173-1
3	2505	82.8	2505	9	US-09-947-199-3
4	2505	82.8	2505	17	US-10-626-173-3
5	2010	66.4	3026	9	US-09-947-199-7
6	2010	66.4	3026	17	US-10-626-173-7
7	1892.2	62.6	2505	9	US-09-947-199-9
8	1892.2	62.6	2505	17	US-10-626-173-9
9	427.6	14.1	740	9	US-09-833-381-1151
10	427.4	14.1	616	9	US-09-833-381-1149
11	414.6	13.7	736	9	US-09-833-381-1150
12	184.2	6.1	304	9	US-09-833-381-1147
13	109.6	3.6	1953	13	US-10-424-599-53404
14	107.4	3.6	3238	13	US-10-424-599-129156

15 100.2 3.3 2207 13 US-10-425-114-22955  
16 100.2 3.3 2327 17 US-10-437-963-54440  
17 98 3.2 1662 9 US-09-938-842A-1014  
18 98 3.2 1662 11 US-09-938-842A-1014  
19 97 3.2 1572 13 US-10-425-114-33402  
20 91.4 3.0 1369 13 US-10-425-114-30542  
21 91.4 3.0 1545 13 US-10-425-114-21691  
22 91.4 3.0 1614 13 US-10-425-114-13674  
23 91.4 3.0 3323 16 US-10-104-047-313  
24 89.4 3.0 972 13 US-10-425-114-6886  
25 88.6 2.9 1146 17 US-10-437-963-59479  
26 88.6 2.9 1461 13 US-10-425-114-4043  
27 87.8 2.9 1219 13 US-10-424-599-130828  
28 87.8 2.9 3020 17 US-10-437-963-65894  
29 87.4 2.9 1299 9 US-09-908-711-12  
30 87.2 2.9 2919 17 US-10-437-963-70898  
31 85.6 2.8 5338 13 US-10-342-887-63  
32 85.6 2.8 5338 13 US-10-172-118-63  
33 85 2.8 1942 13 US-10-425-114-28428  
34 85 2.8 1997 13 US-10-425-114-30767  
35 84.4 2.8 186 9 US-09-833-381-1148  
36 84.2 2.8 5175 9 US-09-964-899-42  
37 84 2.8 799 13 US-10-425-114-30676  
38 84 2.8 14770 10 US-09-873-367C-230  
39 84 2.8 14770 12 US-09-968-007A-124  
40 84 2.8 14770 12 US-09-968-007A-405  
41 84 2.8 14770 13 US-10-342-887-1848  
42 84 2.8 14770 13 US-10-172-118-1848  
43 83.6 2.8 2473 17 US-10-437-963-92781  
44 83.4 2.8 1991 17 US-10-437-963-93604  
45 82.6 2.7 4213 16 US-10-108-260A-794

#### ALIGNMENTS

#### RESULT 1

US-09-947-199-1  
; Sequence 1, Application US/09947199  
; Patent No. US20020127684A1  
; GENERAL INFORMATION:  
; APPLICANT: Raju, Jeyaseelan  
; TITLE OF INVENTION: NOVEL CARC PROTEIN AND NUCLEIC ACID MOLECULES AND USES  
; TITLE OF INVENTION: THEREFOR  
; FILE REFERENCE: WNI-068CE2  
; CURRENT APPLICATION NUMBER: US/09/947,199  
; CURRENT FILING DATE: 2001-09-05  
; PRIOR APPLICATION NUMBER: 60/111,938  
; PRIOR FILING DATE: 1998-12-11  
; PRIOR APPLICATION NUMBER: 09/291,839  
; PRIOR FILING DATE: 1999-04-14  
; PRIOR APPLICATION NUMBER: 09/458,457  
; PRIOR FILING DATE: 1999-12-10  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 3025  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (48)..(2552)  
US-09-947-199-1

Query Match 100.0%; Score 3025; DB 9; Length 3025;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 3025; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTGACCCAGCGTCGGCCCTGGAGAAAGAACTTATATAATGGAATTATA 60

Db 1 GTGACCCAGCGTCGGCCCTGGAGAAAGAACTTATATAATGGAATTATA 60

Qy 61 AATCTAGACCAACCCAACTTGTACTGATGGAAGAAAAAGTCAGTGAATCATATG 120



```
QY 2281 CTTCTGATTGCTGGTGAAACCGGGAGAGACCTGGCGAGTCAATGTCGAGCAATTAAGAA 2340
Db 2281 CTTCTGATTGCTGGTGAAACCGGGAGAGACCTGGCGAGTCAATGTCGAGCAATTAAGAA 2340
QY 2341 GTCGTTTCGAATTCGAATATGCTCTAAATCAAGGTCCTATGCTGCTTTGTCGCCAAGTG 2400
Db 2341 GTCGTTTCGAATTCGAATATGCTCTAAATCAAGGTCCTATGCTGCTTTGTCGCCAAGTG 2400
QY 2401 CTGACAAATATTCCTCTCAAGGTCCTTTTGGAGGAGATGAAAGAAAGTCTTCAATACA 2460
Db 2401 CTGACAAATATTCCTCTCAAGGTCCTTTTGGAGGAGATGAAAGAAAGTCTTCAATACA 2460
QY 2461 CACCCATTGACAAATATGCTATGCTATCCGATCCCATGAGCTCAATGCAATTTCAATCTT 2520
Db 2461 CACCCATTGACAAATATGCTATGCTATCCGATCCCATGAGCTCAATGCAATTTCAATCTT 2520
QY 2521 GCCGAAATAGTAGCAGCTTTTGGAGACAGCAGCTGACAGCAATTCGGGCTATACCTAAGAG 2580
Db 2521 GCCGAAATAGTAGCAGCTTTTGGAGACAGCAGCTGACAGCAATTCGGGCTATACCTAAGAG 2580
QY 2581 AGTTTTTCCCGAATGACAGCAAGCAATTCGAACCAAGCAGCTGGCTTCCAACTATA 2640
Db 2581 AGTTTTTCCCGAATGACAGCAAGCAATTCGAACCAAGCAGCTGGCTTCCAACTATA 2640
QY 2641 ACATTTTACTCTCAAGGTCCTCTTAAATTTGGGCTGTTTTTACTTTGCTCTATTTAATTC 2700
Db 2641 ACATTTTACTCTCAAGGTCCTCTTAAATTTGGGCTGTTTTTACTTTGCTCTATTTAATTC 2700
QY 2701 CCCACTATTAGCAGGCTTTGGATTTGTCCTAAGCAATTAATGCAAAAGAACCAAGACA 2760
Db 2701 CCCACTATTAGCAGGCTTTGGATTTGTCCTAAGCAATTAATGCAAAAGAACCAAGACA 2760
QY 2761 GAATGTATATGAAGATGTTTTTAAATTTGTAATTTAAATTTAAATTTAAATTTAGATCGTTACT 2820
Db 2761 GAATGTATATGAAGATGTTTTTAAATTTGTAATTTAAATTTAAATTTAAATTTAGATCGTTACT 2820
QY 2821 TGGAAATGGAGCCTTAAGTCTGCTGGTGACAGATAATAATATGTTTTCTCGGGCTGAATT 2880
Db 2821 TGGAAATGGAGCCTTAAGTCTGCTGGTGACAGATAATAATATGTTTTCTCGGGCTGAATT 2880
QY 2881 ATGTAGACTTGTGTTGACAGCTATGGGTTTATTTCTTAGAATGTTTCATTTCTTTT 2940
Db 2881 ATGTAGACTTGTGTTGACAGCTATGGGTTTATTTCTTAGAATGTTTCATTTCTTTT 2940
QY 2941 CTCATTATGTTACTCTTAGTGTTCACCTCTGCTGATTAAGATTTCTTTGGTGAATAGAAA 3000
Db 2941 CTCATTATGTTACTCTTAGTGTTCACCTCTGCTGATTAAGATTTCTTTGGTGAATAGAAA 3000
QY 3001 AAAAAAAGGCGGCGC 3025
Db 3001 AAAAAAAGGCGGCGC 3025
```

## RESULT 2

```
US-10-626-173-1
; Sequence 1, Application US/10626173
; Publication No. US20040110232A1
; GENERAL INFORMATION:
; APPLICANT: Raju, Jeyaseelan
; TITLE OF INVENTION: NOVEL CARP PROTEIN AND NUCLEIC ACID MOLECULES AND USES
; FILE REFERENCE: MN1-068CF2
; CURRENT APPLICATION NUMBER: US/10/626,173
; CURRENT FILING DATE: 2003-07-24
; PRIOR APPLICATION NUMBER: US/09/947,199A
; PRIOR FILING DATE: US 60/111,938
; PRIOR APPLICATION NUMBER: US 09/231,839
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: US 09/231,839
; PRIOR FILING DATE: 1999-04-14
; PRIOR APPLICATION NUMBER: US 09/458,457
; PRIOR FILING DATE: 1999-12-10
; NUMBER OF SEQ ID NOS: 9
```

```
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 3025
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (48)..(2552)
US-10-626-173-1
```

```
Query Match 100.0%; Score 3025; DB 17; Length 3025;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3025; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GTGACCCACGCTCCGGCCCTGGAGAAAGAAAGAACTTATAATAATGGGAAATTATA 60
Db 1 GTGACCCACGCTCCGGCCCTGGAGAAAGAAAGAACTTATAATAATGGGAAATTATA 60
QY 61 AATCTAGACCAACCCAAACCTTGTAAGTGAATGGAAGAAAGAAAGTCACTGAATCATATG 120
Db 61 AATCTAGACCAACCCAAACCTTGTAAGTGAATGGAAGAAAGAAAGTCACTGAATCATATG 120
QY 121 TTATCAATATAGAAAGATTAGAAAGTACCTGCAGATCAAGGAAAGAAAGAACTGACAGAAC 180
Db 121 TTATCAATATAGAAAGATTAGAAAGTACCTGCAGATCAAGGAAAGAAAGAACTGACAGAAC 180
QY 181 TAAGGATATATTGGCTCTGATGAAGCCTTCAGTAAAGTCAATTTAAATTTACCGCACTG 240
Db 181 TAAGGATATATTGGCTCTGATGAAGCCTTCAGTAAAGTCAATTTAAATTTACCGCACTG 240
QY 241 AAAATGGGCTGTCTCTACTCTTCAATTTATGTTGCAATTTTGGAGGCAAGAAATCACATATTC 300
Db 241 AAAATGGGCTGTCTCTACTCTTCAATTTATGTTGCAATTTTGGAGGCAAGAAATCACATATTC 300
QY 301 GAACCTCTTATGTTGAAAGGGCTCCGCCATCTCGACTGACAGAAATGGATTTACAGCCT 360
Db 301 GAACCTCTTATGTTGAAAGGGCTCCGCCATCTCGACTGACAGAAATGGATTTACAGCCT 360
QY 361 TGCAATTTAGCAGTTTACAAGGATAATGCAAGATTGATCACTTCTCTGCTTACAGTGAAG 420
Db 361 TGCAATTTAGCAGTTTACAAGGATAATGCAAGATTGATCACTTCTCTGCTTACAGTGAAG 420
QY 421 CTGATATACAGCAGGTTGGATACGGTGGCTCACTGCCCTCCATATTTGTAATATAGCTG 480
Db 421 CTGATATACAGCAGGTTGGATACGGTGGCTCACTGCCCTCCATATTTGTAATATAGCTG 480
QY 481 GCCACCTAGAGGCTGCTGATGCTGTTGCAACATGGAGCTAATGTCATATTTCAAGATG 540
Db 481 GCCACCTAGAGGCTGCTGATGCTGTTGCAACATGGAGCTAATGTCATATTTCAAGATG 540
QY 541 CAGTTTTTTTCACTCCCAATTCGATATTCAGCGTACTATGACATGAACAGGTAACCTCGCC 600
Db 541 CAGTTTTTTTCACTCCCAATTCGATATTCAGCGTACTATGACATGAACAGGTAACCTCGCC 600
QY 601 TTCTTTTGAATTTGGTCTGATGTAATTTAAGTGGTGAAGTTGGAGATAGACCCCTCC 660
Db 601 TTCTTTTGAATTTGGTCTGATGTAATTTAAGTGGTGAAGTTGGAGATAGACCCCTCC 660
QY 661 ACCTAGCATCTGCAAAAGGATTTCTTGAATTTGCAAAACTCTTGATGGAAGAGCAGCA 720
Db 661 ACCTAGCATCTGCAAAAGGATTTCTTGAATTTGCAAAACTCTTGATGGAAGAGCAGCA 720
QY 721 AAGCAGATGTAATGCTCAAGATAATGAAGACCATGTCCTCCACTCCATTTCTGTTCCGAT 780
Db 721 AAGCAGATGTAATGCTCAAGATAATGAAGACCATGTCCTCCACTCCATTTCTGTTCCGAT 780
QY 781 TTGACACCATGATATAGTTAAGTATCTGCTCAAAAGGATTTGGAAGTTCAACCTCATG 840
Db 781 TTGACACCATGATATAGTTAAGTATCTGCTCAAAAGGATTTGGAAGTTCAACCTCATG 840
QY 841 TTGTTAATATCTATGGAGATACCCCTTACACCTGGCATGCTACAAATGGCAAAATTTGAAG 900
Db 841 TTGTTAATATCTATGGAGATACCCCTTACACCTGGCATGCTACAAATGGCAAAATTTGAAG 900
```

[illegible]

```

RESULT 3
US-09-947-199-3
; Sequence 3, Application US/09947199
; Patent No. US20020127684A1
; GENERAL INFORMATION:
; APPLICANT: Raju, Jeyaseelan
; TITLE OF INVENTION: NOVEL CARK PROTEIN AND NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: MNI-068CP2
; CURRENT APPLICATION NUMBER: US/09/947,199
; PRIOR FILING DATE: 2001-09-05
; PRIOR APPLICATION NUMBER: 60/111,938
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/291,839
; PRIOR FILING DATE: 1999-04-14
; PRIOR APPLICATION NUMBER: 09/458,457
; PRIOR FILING DATE: 1999-12-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2505
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2505)
US-09-947-199-3

Query Match      82.8%; Score 2505; DB 9; Length 2505;
Best Local Similarity 100.0%; Pred.No. 0;
Matches 2505; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY      48  ATGGGAATTATAAATCTAGACCAACCCAAACTTGTACTGATCAATGGAGAAAAAGTC 107
DB      1  ATGGGAATTATAAATCTAGACCAACCCAACTTGTACTGATCAATGGAGAAAAAGTC 60

QY     108  AGTGAATCATATGTTTATCACATGAAGAATTAGAAGTACCTCGAGATCAAGGAAAAA 167
DB     61  AGTGAATCATATGTTTATCACATGAAGAATTAGAAGTACCTCGAGATCAAGGAAAAA 120

QY     168  GAACTGACAGAACTAAGGAATATATTTGGCTCTCATGAAGSCCTTCAGTAAAGTCAATTTA 227
DB     121  GAACTGACAGAACTAAGGAATATATTTGGCTCTCATGAAGSCCTTCAGTAAAGTCAATTTA 180

QY     228  AATTACCGCACTGAAAATGGGCTGCTCTACTCTCATTTATTTGTGCATTTGTGGAGCAAG 287
DB     181  AATTACCGCACTGAAAATGGGCTGCTCTACTCTCATTTATTTGTGCATTTGTGGAGCAAG 240

QY     288  AAATCACATATTCGAACCTTTATGTTGAAGGCGCTCGGCCATCTCGACTGACAGAAT 347
DB     241  AAATCACATATTCGAACCTTTATGTTGAAGGCGCTCGGCCATCTCGACTGACAGAAT 300

QY     348  GGATTTACAGCCCTTGCATTTAGCAGTTTACAAGGATAATGCAGAAATGATCACTTCTCTG 407
DB     301  GGATTTACAGCCCTTGCATTTAGCAGTTTACAAGGATAATGCAGAAATGATCACTTCTCTG 360

QY     408  CTTTCAGTGGAGCTGATATACAGAGTTGGATACGGTGGCCTCTACTGCGCTTCCATATT 467
DB     361  CTTTCAGTGGAGCTGATATACAGAGTTGGATACGGTGGCCTCTACTGCGCTTCCATATT 420

QY     468  GCTACAAATAGCTGGCCACCTAGAGGCTGCTGATGCTGTTGCAACATCGAGCTAATGTC 527
DB     421  GCTACAAATAGCTGGCCACCTAGAGGCTGCTGATGCTGTTGCAACATCGAGCTAATGTC 480

QY     528  AATATTCAAGATGCAAGTTTTTTTCACTCCATTGCATATTCACGCTACTATGGACATGAA 587
DB     481  AATATTCAAGATGCAAGTTTTTTTCACTCCATTGCATATTCACGCTACTATGGACATGAA 540

QY     588  CAGGTAACTCGCGCTTCTTTTGAATTTGGTGTCTGATGTAATGTAAGTGGTGAAGTTGGA 647
DB     541  CAGGTAACTCGCGCTTCTTTTGAATTTGGTGTCTGATGTAATGTAAGTGGTGAAGTTGGA 600

QY     648  GATAGACCCCTCCACCTAGCATCTGCAAAAAGGATTCTTGAATATTCAAAACTCTTTGATG 707

```

Db	1681	CAGTCTAAATTAAATTATTTGAGTAGATGTTGGCAAGGCATGGAGTACCTTCAACAACCTG	1741
Qy	1788	ACACAGCAAAATTATACATCGTGAATCTGAAACAGTCACAATATTTCTTCTATGAGGATGGG	1847
Db	1741	ACACAGCAAAATTATACATCGTGAATCTGAAACAGTCACAATATTTCTTCTATGAGGATGGG	1800
Qy	1848	CATGCTGTGGTGGCAGATTTTGGAGAAATCAAGATTTCTACAGTCTCTGGATGACAGACAAC	1907
Db	1801	CATGCTGTGGTGGCAGATTTTGGAGAAATCAAGATTTCTACAGTCTCTGGATGACAGACAAC	1860
Qy	1908	ATGACAAAACCACTGGGAACCTCCGTTGGATGGCTCTGAGGTGTTCAACGAGTGCAC	1967
Db	1861	ATGACAAAACCACTGGGAACCTCCGTTGGATGGCTCTGAGGTGTTCAACGAGTGCAC	1920
Qy	1968	CGGTACACCATCAAGACAGATGTTCTCAGCTATGCTCTGTGTCTGTGGAAATTTCTCACT	2027
Db	1921	CGGTACACCATCAAGACAGATGTTCTCAGCTATGCTCTGTGTCTGTGGAAATTTCTCACT	1980
Qy	2028	GGCGAAATTCATTCGCTCATCTCAAGCAGCGGCTGCGGACGACAGACATGGCTTACCA	2087
Db	1981	GGCGAAATTCATTCGCTCATCTCAAGCAGCGGCTGCGGACGACAGACATGGCTTACCA	2040
Qy	2088	CACATCAGACCTCCCATTTGGCTATTTCATTTCCCAAGCCCATCATCTCTGCTGATACGA	2147
Db	2041	CACATCAGACCTCCCATTTGGCTATTTCATTTCCCAAGCCCATCATCTCTGCTGATACGA	2100
Qy	2148	GGGTGGAAACGCAATGCTCGTAAGGAAGACCCGAATTTTCTTGAAAGTTGTGATGAAGTTAGAA	2207
Db	2101	GGGTGGAAACGCAATGCTCGTAAGGAAGACCCGAATTTTCTTGAAAGTTGTGATGAAGTTAGAA	2160
Qy	2208	GAGTGTCTCTGCACATTTAGCTGATGTCCTCGCATCAAGTAAACAGACAGTGGGTCTCTC	2267
Db	2161	GAGTGTCTCTGCACATTTAGCTGATGTCCTCGCATCAAGTAAACAGACAGTGGGTCTCTC	2220
Qy	2268	TCACCTCTCTTCTTCTTGATTGCTGGTGAACCGGGGAGGACCTGGCGGAGTCAATGTG	2327
Db	2221	TCACCTCTCTTCTTCTTGATTGCTGGTGAACCGGGGAGGACCTGGCGGAGTCAATGTG	2280
Qy	2328	GCAGCATTGAAGAGTCGTTTCGAATTCGAATATGCTCTTAAATGCAAGGTCTCTATGCTGCT	2387
Db	2281	GCAGCATTGAAGAGTCGTTTCGAATTCGAATATGCTCTTAAATGCAAGGTCTCTATGCTGCT	2340
Qy	2388	TTGTCCCAAGTGTGGACAATATTCCTCTCAAGGTCTGCTTTGGAGGAGATGAAAAGA	2447
Db	2341	TTGTCCCAAGTGTGGACAATATTCCTCTCAAGGTCTGCTTTGGAGGAGATGAAAAGA	2400
Qy	2448	AGTCTTCAATACACACCCATTGCAAAATATGGCTATGATCCGATCCCATGAGCTCAATG	2507
Db	2401	AGTCTTCAATACACACCCATTGCAAAATATGGCTATGATCCGATCCCATGAGCTCAATG	2460
Qy	2508	CAATTTTCATCTTGGCGAAATAGTAGCAGCTTTTGAGGACAGCAGC	2552
Db	2461	CAATTTTCATCTTGGCGAAATAGTAGCAGCTTTTGAGGACAGCAGC	2505

601	GATAGACCCCTCCACTAGCATCTGCAAAAGGATTCCTGAATATTCGAAAACCTCTTGATG	DB
648	GATAGACCCCTCCACTAGCATCTGCAAAAGGATTCCTGAATATTCGAAAACCTCTTGATG	QY
708	GARGAAGGACGACAAAGCAGTGTGAATGCTCAAGATAATGAAGACCATGTCCACCTCCAT	QY
661	GARGAAGGACGACAAAGCAGTGTGAATGCTCAAGATAATGAAGACCATGTCCACCTCCAT	DB
768	TTCCTGTTCTCGATTTGGACACATGATATAGTTAAGTATCTGCTGCAAAAGTGATTTGGAA	QY
721	TTCCTGTTCTCGATTTGGACACATGATATAGTTAAGTATCTGCTGCAAAAGTGATTTGGAA	DB
828	GTTTCAACCTCATGTTGTTAAATATCTATGGAGATACCCCTTACACCTGGCATGCTACAAT	QY
781	GTTTCAACCTCATGTTGTTAAATCTATGGAGATACCCCTTACACCTGGCATGCTACAAT	DB
888	GGCAAAATTTGAAGTTGCCAAGGAATCATCCAAATATCAGGAACAGAAAGCTCTGACTAAG	QY

Db 841 GGCAAAATTGAGTTGCCAAGGAAATCATCCAAATATCAGGAACAGAAAGTCTGACTAAG 900  
Qy 948 GAAACAATCTTTCAGTGAACACAGCTTTTCATAGTGTCTGTACCTATGCGCAAGAGCATTTGAC 1007  
Db 901 GAAACAATCTTTCAGTGAACACAGCTTTTCATAGTGTCTGTACCTATGCGCAAGAGCATTTGAC 960  
Qy 1008 TAGTCAAAATTTCTTTCATGATCAGAAATGTCTATAAACAATCAACACCAAGGAGGATGG 1067  
Db 961 TAGTCAAAATTTCTTTCATGATCAGAAATGTCTATAAACAATCAACACCAAGGAGGATGG 1020  
Qy 1068 CACACTGATTACACTCTGCTGTGTACACGGTTCACATTCGCTGTGTTCAGTTCCTTACTG 1127  
Db 1021 CACACTGGAATTACACTCTGCTGTGTACACGGTTCACATTCGCTGTGTTCAGTTCCTTACTG 1080  
Qy 1128 GATTAATGGAGCTGATGAATTCATAGTGGCTTGTATCCAGCAGGTCCTAGTGGTGA AAAA 1187  
Db 1081 GATTAATGGAGCTGATGAATTCATAGTGGCTTGTATCCAGCAGGTCCTAGTGGTGA AAAA 1140  
Qy 1188 GATCAGCAGACATGTTTGTATGTGGCTTATGAAAAGGGCATGATGCCATTTGTCACTC 1247  
Db 1141 GATCAGCAGACATGTTTGTATGTGGCTTATGAAAAGGGCATGATGCCATTTGTCACTC 1200  
Qy 1248 CTGAAGCATTATAAGACACACAGATGAATTCCTGTGAATGAATATTTCTCAGCCTGGA 1307  
Db 1201 CTGAAGCATTATAAGACACACAGATGAATTCCTGTGAATGAATATTTCTCAGCCTGGA 1260  
Qy 1308 GGAGATGGCTCCTATGTCTGTCTGTTCATCACCTTGGGGAGATTAAAAGCATGACAAA 1367  
Db 1261 GGAGATGGCTCCTATGTCTGTTCATCACCTTGGGGAGATTAAAAGCATGACAAA 1320  
Qy 1368 GAGAAGGCAGATATTTCTCCTTAAGAGCTGGATGGCTTTCACATTTCCATCTTCAGCTC 1427  
Db 1321 GAGAAGGCAGATATTTCTCCTTAAGAGCTGGATGGCTTTCACATTTCCATCTTCAGCTC 1380  
Qy 1428 TCAGAAATTCAGTTCCTATGAGATTTGGCTCAGGTTCTTTGGGAAAGTATATAAGGA 1487  
Db 1381 TCAGAAATTCAGTTCCTATGAGATTTGGCTCAGGTTCTTTGGGAAAGTATATAAGGA 1440  
Qy 1488 CGATGCAGAAATAAATAGTGGCTATAAAGCTTATCGAGCCAAATACCTACTGCTCCAA 1547  
Db 1441 CGATGCAGAAATAAATAGTGGCTATAAAGCTTATCGAGCCAAATACCTACTGCTCCAA 1500  
Qy 1548 TCAGATGTGATATGTTTTCGAGAGGTTCCATTTCTGCGAGCTCAATCATCCCTGC 1607  
Db 1501 TCAGATGTGATATGTTTTCGAGAGGTTCCATTTCTGCGAGCTCAATCATCCCTGC 1560  
Qy 1608 GTAAATTCAGTTTGGGTGCTTCTGAAATGATCCAGCCAGCTTGGCCATTTGCTACTCA 1667  
Db 1561 GTAAATTCAGTTTGGGTGCTTCTGAAATGATCCAGCCAGCTTGGCCATTTGCTACTCA 1620  
Qy 1668 TACATATCAGGGGTTCTCTGTTCTCCCTCTTCATGACAGAGAGGATTTCTGATTTG 1727  
Db 1621 TACATATCAGGGGTTCTCTGTTCTCCCTCTTCATGACAGAGAGGATTTCTGATTTG 1680  
Qy 1728 CAGTCTAAATTAATATTCAGTAGATGTTGCCAAAGGAGTGGAGTTCACAACTG 1787  
Db 1681 CAGTCTAAATTAATATTCAGTAGATGTTGCCAAAGGAGTGGAGTTCACAACTG 1740  
Qy 1788 ACACAGCAATTAATACATCTGACTTGAACAGTCACAATATTTCTCTATCAGAGATGG 1847  
Db 1741 ACACAGCAATTAATACATCTGACTTGAACAGTCACAATATTTCTCTATCAGAGATGG 1800  
Qy 1848 CATGCTGTGTGGCAGATTTTGAGAAATCAAGATTTCTACGTCTCTGGATGAAGACAAC 1907  
Db 1801 CATGCTGTGTGGCAGATTTTGAGAAATCAAGATTTCTACGTCTCTGGATGAAGACAAC 1860  
Qy 1908 ATGACAAAACAACCTGGACCTCCGTTGGATGGCTCCGTTGGATGGCTTTCACAGTGCAC 1967  
Db 1861 ATGACAAAACAACCTGGACCTCCGTTGGATGGCTCCGTTGGATGGCTTTCACAGTGCAC 1920  
Qy 1968 CGGTACACCATCAAGCAGATGCTTTCAGCTATGCTGTGTGTGTGGGAAATTTCTCACT 2027  
Db 1921 CGGTACACCATCAAGCAGATGCTTTCAGCTATGCTGTGTGTGTGGGAAATTTCTCACT 1980

Qy 2028 GCGGAAATTCATTGCTCATCTCAAGCCAGCGGCTCGCGCAGCAGACATGGCTTACCAC 2087  
Db 1981 GCGGAAATTCATTGCTCATCTCAAGCCAGCGGCTCGCGCAGCAGACATGGCTTACCAC 2040  
Qy 2088 CAGATCAGACTCCCATTTGCTTATTCATCCCAAGCCCATATCATCTCTGCTGATACA 2147  
Db 2041 CAGATCAGACTCCCATTTGCTTATTCATCCCAAGCCCATATCATCTCTGCTGATACA 2100  
Qy 2148 GGGTGGAAACGATGCTCTGAAGGAAGACCCGAATTTTCTGAAGTTGTCATGAAGTTAGAA 2207  
Db 2101 GGGTGGAAACGATGCTCTGAAGGAAGACCCGAATTTTCTGAAGTTGTCATGAAGTTAGAA 2160  
Qy 2208 GAGTGTCTCTGCAACATTCAGTGTCTCTGCAATCAAGTAAACAGAGTGGGTCTCTC 2267  
Db 2161 GAGTGTCTCTGCAACATTCAGTGTCTCTGCAATCAAGTAAACAGAGTGGGTCTCTC 2220  
Qy 2268 TCACCTTCTTCTTCTTCTGATTCGCTGGTGAACCGGGAGGACCTGGCCGGAGTCATGT 2327  
Db 2221 TCACCTTCTTCTTCTTCTGATTCGCTGGTGAACCGGGAGGACCTGGCCGGAGTCATGT 2280  
Qy 2328 GCAGCATTAAGAAGTGGTTTTCGAATTTGGAATATGCTCTAAATGCAAGGTCCTATGCTGT 2387  
Db 2281 GCAGCATTAAGAAGTGGTTTTCGAATTTGGAATATGCTCTAAATGCAAGGTCCTATGCTGT 2340  
Qy 2388 TTGTCCCAAAAGTGGTGGACAAATATTCCTCTCAAGTCTCTTTGGAGGAGATGAAGA 2447  
Db 2341 TTGTCCCAAAAGTGGTGGACAAATATTCCTCTCAAGTCTCTTTGGAGGAGATGAAGA 2400  
Qy 2448 AGTCTTCAATACACACCCATTGACAAATATGCTATGATATCGATCCCATGAGCTCAATG 2507  
Db 2401 AGTCTTCAATACACACCCATTGACAAATATGCTATGATATCGATCCCATGAGCTCAATG 2460  
Qy 2508 CATTTTCATTTCTTGGCCGAATATAGTAGCAGCTTTTGAGGACAGCAGC 2552  
Db 2461 CATTTTCATTTCTTGGCCGAATATAGTAGCAGCTTTTGAGGACAGCAGC 2505

RESULT 5  
US-09-947-199-7  
; Sequence 7, Application US/09947199  
; Patent No. US20020127684A1  
; GENERAL INFORMATION:  
; APPLICANT: Raju, Jeyaseelan  
; TITLE OF INVENTION: NOVEL CARP PROTEIN AND NUCLEIC ACID MOLECULES AND USES  
; TITLE OF INVENTION: THEREFOR  
; FILE REFERENCE: MNI-068CP2  
; CURRENT APPLICATION NUMBER: US/09/947,199  
; PRIOR FILING DATE: 2001-09-05  
; PRIOR APPLICATION NUMBER: 60/111,938  
; PRIOR FILING DATE: 1998-12-11  
; PRIOR APPLICATION NUMBER: 09/291,839  
; PRIOR FILING DATE: 1999-04-14  
; PRIOR APPLICATION NUMBER: 09/458,457  
; PRIOR FILING DATE: 1999-12-10  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 7  
; LENGTH: 3026  
; TYPE: DNA  
; ORGANISM: Rattus norvegicus  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (61)..(2565)  
US-09-947-199-7

Query Match 66.4%; Score 2010; DB 9; Length 3026;  
Best Local Similarity .80.8%; Pred. No. 0;  
Matches 2425; Conservative 0; Mismatches 570; Indels 13; Gaps 6;  
Qy 16 CGGCCCTGGAGAAAGGAAGAACTTATAATAATGGAAATTTATAAATCTAGACCAACCC 75  
Db 29 CAGCAGGAGAGAAAGCAAGAGCTTCTTTAAATGGGAAATTACAAATCCAGACCAACAC 88

QY	1156	CTTGTGATCCAGCAGCTCTAGTGGTGAAGATGAGCAGACATGTTTGTGTGGCTT	1215
Db	1169	CTTGTGATCCAGCAGCTCTAGTGGTGAAGATGAGCAGACATGTTTGTGTGGCTT	1228
QY	1216	ATGAAAAAGGCGATGATGCTTGTACACTCTCTGAAGCATTTATAGAGACACCAAGATG	1275
Db	1229	ACGAGAAAGGCGATGATGCTTGTACACTCTCTGAAGCATTTATAGAGACACCAAGATG	1288
QY	1276	AATTGCCCTGTATGAAATTTCTCAGCTGGAGGAGATGGCTCTCTATGCTGTGTTCCAT	1335
Db	1289	AGCTGCCATGTAAGCAATTTCTCCAGCTGGAGGAGATGGCTCTCTATGCTGTGTTCCAT	1348
QY	1336	CACCTTTGGGAGAGATTAAGCATGACAAAAGAGAGGAGAGATTTCTCTCTCTAAAGAG	1395
Db	1349	CCCTTTGGGCAAGATTAAGCATGACAAAAGAGAGGAGAGATTTCTCTCTCTGAGG	1408
QY	1396	CTGATTCGCTTACATTTCCATCTTCAGCTCTCAGAAATTCAGTTCCATGAGATATTG	1455
Db	1409	CTGAATACCTCTCCGCTTCCATCTCCAACTCTCCGAAATTCGAGTTCCAGGATATTG	1468
QY	1456	GCTCAGGTTCTTTTGGGAAAGTATATAAGGACGATGCAGAAATAAATAGTGGCTATAA	1515
Db	1469	GCTCGGTTCTTTTGGGAAAGTCTATAAGGCGCATGCAGAAATAAATAGTGGCGATCA	1528
QY	1516	AACGTTATCGAGCCAAATACCTACTGCTCCAAAGTCAGATGCGAATATGTTTTCGCGAGG	1575
Db	1529	AACGATACCGAGCCAAACACCTACTGCTCCAAAGTCAGACGCTGGATATGTTTTCGCGAGG	1588
QY	1576	TGTCATTTCTGCGAGCTCAATCATCTCCGCTTAATTCAGTTTGGGCTGCTTGTCTGA	1635
Db	1589	TGTCATTTCTGCGAGCTCAACCACTCCGCTGCTTAATTCAGTTTGGGCTGCTTGTCTGA	1648
QY	1636	ATGATCCAGCCAGTTTGGCAATTTGCTCACTCAATACATATACAGGGGCTTCTCTGTTCTCC	1695
Db	1649	ATGATCCAGCCAGTTTGGCAATTTGCTCACTCAATACATATACAGGGGCTTCTCTGTTCTCC	1708
QY	1696	TCCTTCATGAGCAGAGAGATTTCTGATTTGCACTTAATTAATTAATTCAGTAGATG	1755
Db	1709	TGCTTCATGAACAGAGAGATTTCTGATTTGCACTTAATTAATTAATTCAGTAGATG	1768
QY	1756	TTGCAAGGCGATGAGTACTCTCAAACTGACACAGCAATATATACATCTGATCTGA	1815
Db	1769	TTGCAAGGCGATGAGTACTCTCAAACTGACACAGCAATATATACATCTGATCTGA	1828
QY	1816	ACAGTCAATATTTCTTCTATAGAGATGGGATGCTGTGTTGGGAGATTTTGGAGAT	1875
Db	1829	ACAGTCAATATTTCTTCTATAGAGATGGGATGCTGTGTTGGGAGATTTTGGAGAT	1888
QY	1876	CAAGATTTCTACAGTCTCTGATGAAGACAAACATGACAAACCTGGGAACCTCCGTT	1935
Db	1889	CAAGATTTCTGATGCTCTGATGAAGACAAACATGACAAACCTGGGAACCTCCGTT	1948
QY	1936	GGATGGCTCTGAGGTTTCAAGCAGTGCATCGGTACACCATCAAAAGCAGATGTTTCA	1995
Db	1949	GGATGGCTCTGAGGTTTCAAGCAGTGCATCGGTACACCATCAAAAGCAGATGTTTCA	2008
QY	1996	GCTATGCTGTGCTCTGTTGGGAAATTTCTCAGCTGGGAAATTTCCATTCGCTCATCAAG	2055
Db	2009	GTTACTCCCTGTGCTGTGGGAGCTCTCAGCTGGGAAATTTCCATTCGCTCATCAAG	2068
QY	2056	CAGCGCTGCGCAGCAGACATGGCTTACCACCATCAGACCTCCCATTTGGCTATTCCA	2115
Db	2069	CAGCGCTGCGCAGCAGACATGGCTTACCACCATCAGACCTCCCATTTGGCTATTCCA	2128
QY	2116	TTCCAGCCCATATCTCTGCTGATAGAGGTTGGAAGCAGATGCTCTGAGGAGAC	2175
Db	2129	TTCCAGCCCATATCTCTGCTGATAGAGGTTGGAAGCAGATGCTCTGAGGAGAC	2188
QY	2176	CGAATTTCTGAAATTTGTCATGAAGTTAGAAGTGTCTCTGCAACATTTGAGCTGATGT	2235
Db	2189	CAGATTTCTGAAATTTGTCATGAAGTTAGAAGTGTCTCTGCAACATTTGAGCTGATGT	2248
QY	2236	CTCTGCTCAAGTAAACAGCAGTGGGCTCTCTCAGCCTTCTTCTTCTGATGCTGCG	2295

Db	2249	CTCCAGCATCAAGTAACAGCAGTGGCTCTCTGTGACCTTCTCTTCTCCGATTGCGCTGC	2308
Qy	2296	TGAACCGGGGAGGACCTGGCGGAGTCATGTGCAGCATTAAGAAAGTCGTTTCGAATTGG	2355
Db	2309	TCAGCCGGGAGGCGCTGGCGGAGCCAGCTGCAGCCCTTACGAGCCGTTTTCGAGTTGG	2368
Qy	2356	AATATGCTCTAAATGCAAGTCTTATGCTGCTTTGTGCCAAGTCTGTGACAATATTCCT	2415
Db	2369	AGTATGCCCTTAAATGCAAGTCTTATGCTGGGTGGTCCCAAAAGTTGTGGAAACACACTCTA	2428
Qy	2416	CTCAAGGTCTGTCTTTGGAGGAGATGAAAAGAGTCTTCAATACACACCCATTCACAAT	2475
Db	2429	ATCCGGGCGCTGCTTTGGAGGAGATGAATAGGAGCACCCAGTATTCAACTGTTGACAAAT	2488
Qy	2476	ATGGCTATGTATCCGATCCCATGAGCTCAATGCAATTTCAITCTTTCGCGAAATAGTAGCA	2535
Db	2489	ACGGGTATGTGTCGTGATCCCATGAGCCTGACGACACCTTCACCTCCGCGCAAGACGACAGCA	2548
Qy	2536	GCTTTGAGGACAGCAGCTGACAGCATTTGGCGGTATACCTTAAGGACAGTTTTTTCCCGCAA	2595
Db	2549	ACTTTGAGGACAGCAACTGACAG-GTCTGGCATACACTTAAGGGCGTCTCCCATCAGG	2607
Qy	2596	CTGACAGCAACGATTCGAAACGAGCAAGCTGGCTTCCAACTATAACATTTTACTCTCAA	2655
Db	2608	CTGACAGCAGTGAATTTTACCCTATGGCAGGCTTGCTTTCCAAATATTAACGCCCTGCGCTCTG	2667
Qy	2656	AGGTCCTCCTAAATGGGCTTGTTTTTACTGTGCTTATTTAATCCCCACATTTAGCAGG	2715
Db	2668	AGGT-TTCTTTCAAAATCGTCTTGCTTAATCTAAGCTCGTTTTAATCCCTTCTACAGGACAG	2726
Qy	2716	CTTTGGATTTGTGCTAAGGAATAATATGCAAAAGAACCAAGACAGAAATGTATATGAAGA	2775
Db	2727	GCTTTGACTCATGCC--AAGCGTAAGTGTCAAAGAGCAGATACAGAATGTGCATGAGGA	2784
Qy	2776	ATTGTTTTTATTTTGTAAATTAATAAAAAATTTAGATCGTTACTTGTGAATATGAGGCTTA	2835
Db	2785	ATTGTTCTTAGTTTGAATTTAAAGCGCCTTAATTCGCTGGGGCTGGGGTTCAAA---TC	2840
Qy	2836	AGTCTGTGGTGGACAGATAAATATATGTTTTTCTGGGCTGAATTTATGTAGACTTGTGTT	2895
Db	2841	TGTGTAGATAGCTGGGTGACCCCTATGATTTGTAGACAAACGTGTGGGCTTGTGTT	2900
Qy	2896	TGACAG-CTATGGGTTTATTTCTTAGACAATTTGTTCAATTTCTTTTCTCATATGTTACT	2954
Db	2901	TGAGGGTCTCGTGTGGGTTTCTTAAAAACAAGCTGGCTGATTTATCTCTCTGTTGCGCTTT	2960
Qy	2955	TCTAGTGTTCACCTCTGTGATTAAGATTTCTTTGGTGAATAAGAAAAAATAAAAAA	3014
Db	2961	----GTTGTTACTTCTGTGATTAAGTCTCTCGGTGATCTAGAAAAAATAAAAAA	3016
Qy	3015	AGGGCGC	3022
b	3017	GGGCGCC	3024

## RESULT 6

RESULT 6  
 US-10-626-173-7  
 ; Sequence 7, Application US/10626173  
 ; Publication No. US20040110232A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Raju, Jeyaseelan  
 ; TITLE OF INVENTION: NOVEL CARX PROTEIN AND NUCLEIC ACID MOLECULES AND USES  
 ; TITLE OF INVENTION: THEREFOR  
 ; FILE REFERENCE: MNI-068CP2  
 ; CURRENT APPLICATION NUMBER: US/10/626,173  
 ; CURRENT FILING DATE: 2003-07-24  
 ; PRIOR APPLICATION NUMBER: US/09/947,199A  
 ; PRIOR APPLICATION NUMBER: US 60/111,938  
 ; PRIOR FILING DATE: 1998-12-11  
 ; PRIOR APPLICATION NUMBER: US 09/291,839  
 ; PRIOR FILING DATE: 1999-04-14  
 ; PRIOR APPLICATION NUMBER: US 09/458,457

Db	869	GTGACACTCCTTTGACCTGGCATCTACAAATGGAATTTTGAAGTTGGCAAGAAATTTG	928	1949	GGATGGCCCTGAGGTGTTTCAACAGTGCACGAGATACACCATCAAGGCTGATGCTTCA	2008
Qy	916	TCCAAATATCAGGAACAGAAAGTCTGACTAAGGAAACATCTTTCAAGTAAACAGCTTTTC	975	1996	GCTATGCTGTGTCTGTGTGGGAAATTTCTCACTGGGGAAATTTCCATTCGCTCATCTCAAGC	2055
Db	929	TCCAGGTAAAGAACTGAAAGTCTGACTAAGGAAACATCTTCAGGAGACAGCTTTTC	988	2009	GTTACTCCCTGTGTGTGTGGAGCTCCTCACTGGGAAATTTCCATTCGCTCATCTCAAGC	2068
Qy	976	ATAGTCTTGTACCTATGGCAAGACATTTGACCTAGTCAAAATTTCTTTGTATGACAGATG	1035	2056	CAGCGCTGGGAGAGACAGATGGCTTACCACATCAGACCTCCCATTTGGCTATTCCA	2115
Db	989	ACAGTGTGTACCTATGGCAAGAACATTTGACCTGGTCAAAATTTCTTTGTATGACAGATG	1048	2069	CAGCGCTGGAGAGAGATATGGCTATCAACATCAGACCGGCCATCGGCTATTCCA	2128
Qy	1036	TCATTAACATCAACACCAAGAGGATGGGACACTGGATTTACACTCTGCTTGTACCTACC	1095	2116	TTCCCAAGCCCATATCATCTCTGATACGAGGTGGAAACCATGTCTCTGAAGAGAC	2175
Db	1049	CTGTGAACATTAACACCGAAGAGATGGGACACAGGATTTGCACTCTGCTTGTACCTACC	1108	2129	TCCCAAGCCCATCTCATCTCTGATACGCGGCTGGAATTCATGTCTCTGAAGACGAC	2188
Qy	1096	ACGGTCAATTCGGCTGGTTCAGTTCTTATGGATATGGAGCTGATATGAATCTAGTGG	1155	2176	CCGAAATTTCTGAAGTTGTCAATGAATTTAGAGAGTGTCTCTGCAACATTTGAGCTGATGT	2235
Db	1109	ACGGCCATATCCGCTGGTTCAGTTCTTACTTGATATGCTGCAGATATGAATCTTGTGG	1168	2189	CAGAGTTCTCTGAAGTCGTTAGCAAACTGGAGAGTGCCCTATGCAATGTGAGCTCATGT	2248
Qy	1156	CTTGTGATCCAGGAGTCTAGTGGTGAAGAAAGATGAGACATGTTTGTATGTTGGCTT	1215	2236	CTCTGTCATCAAGTAAACAGCAGTGGGTCTCTCTCACCTTTCTTCTTCTGATTGCCCTGG	2295
Db	1169	CTTGTGATCCAGCAGTCTAGTGGTGAAGAAAGATGAGACATGTTTGTATGTTGGCTT	1228	2249	CTCAGCATCAAGTAAACAGCAGTGGGTCTCTCTCACCTTTCTTCTTCCGATTCCTGTC	2308
Qy	1216	ATGAAAAAGGCGATGCCATTTGTCACTCTCTGAAGCATTTATAAGAGACCAACAAGATG	1275	2296	TGAACCGGGAGGAGACCTGGCCGAGTCTATGTGGCAGCATTAAGAGATGCTTTCGAAATGG	2355
Db	1229	ACGAGAAAGGACATGATGCCATTTGTACACTCTCTGAAGCATTTACACTCTGAGACACTACNAGAGACCCAGGAGG	1288	2309	TGAGCCGGGAGGCGCTGGCCGAGCCAGCTGGCAGCCCTTACGGAGCCGTTTTTGAGTTGG	2368
Qy	1276	AATGCCCTGTAATTAATTTCTCAGCTCTGAGAGATGGCTCTTATGTGTCTGTTCCAT	1335	2356	AATATGCTCTAAATGCAAGTCTCTATGCTGCTTTGTCCCAAGTGTCTGCAATATTCTCT	2415
Db	1289	AGCTGCCATGTAAGCAATTTCCAGCCTGGAGGAGTGGCTCTTATGTGTCTGTTCTT	1348	2369	AGTATGCCCTTAATGCAAGTCTCTATGCTGGGTGGTCCCAAGTGTGGAAACACACTCTA	2428
Qy	1336	CACCTCTGGGAGATTAAGAGCATGACAAAGAGAGGAGCATTTCTCTCTTAAGAG	1395	2416	CTCAAGTCTGTCTTTGGAGGAGATGAAAGAGTCTTCAATACACACCCATTGACAAAT	2475
Db	1349	CCCCCTTGGGCAAGATTAAGAGCATGACAAAGAGAGGAGATGTTCTCTCTCCTGAGGG	1408	2429	ATCGGCTGTCTTTGGAGGAGATGAATAGGAGCACCCAGTATTCAACTGTGTGACAAAT	2488
Qy	1396	CTGATTCCTTCATTTCCATCTTCCAGCTCTCAGAAATTTGAGTTCCATGAGATTATTG	1455	2476	ATGGCTATGATCCGATCCCATGAGCTCAATGCATTTTCATTTCTTCCGCAATATAGCA	2535
Db	1409	CTGAACACTCCCTCCGCTTCCATCTCCAACTCTCCGAAATCGAGTTCCACGAGATTATCG	1468	2489	ACGGCTATGCTGTGATCCCATGAGCTGACGACCTTCACTCCGCCCAAGACGACGCA	2548
Qy	1456	GCTCAGTCTTTTGGGAAAGTATATAAGACGATCGCAAAATTAAGATGGCTATAA	1515	2536	GCTTTGAGGACAGACGCTGACAGATTCCGCGTATACCTTAAGAGAGTTTTTTCCCCGAA	2595
Db	1469	GCTCGGTTCTTTTGGGAAAGTCTATAAGAGGCGATGCAAAATTAAGATGGCGATCA	1528	2549	ACTTTGAGGACAGCAACTGACAG-GTCTGGCATACACTTAAGGGCGTCTCCCCATCAGG	2607
Qy	1516	AACGTTATCGAGCCAAATACCTACTGCTCCAACTGTCAGATGTGGATGTTTTCGCGAGAG	1575	2596	CTGACAGCAAGATTCCAACCGCAAGTGGCTGCCAACTATACATTTTACTCTCAA	2655
Db	1529	AACGATACGAGCCAAACCTACTGCTCCAACTGTCAGATGTGGATGTTTTCGCGAGAGG	1588	2608	CTGACAGCAGTATTTTACCCATGGCAGGCTTGCCTTCCAAATTAACGCCCTGCCCTCTG	2667
Qy	1576	TGTCATCTCTGCAGCTCAATCATCTCCGCTGAATTCAGTTTGTGGTGTCTGCTTGA	1635	2656	AGGTCTCCTTAAATTTGGGCTTGTCTTACTTGTCTTATTTAATTTCCCACTATTAGCAGG	2715
Db	1589	TGTCATCTCTGCAGCTCAACACCGCTGCGTGTGATGTTTGTGGTGTCTGCTGCTGG	1648	2668	AGGT-TTCTTCAATCGTCTGCTTATTCTAAGCTCGTTTAAATTCCTTCTACAGACAG	2726
Qy	1636	ATGATCCAGCCAGTTTCCCATTTGCTCAATACATATCAGGGGTTCTCTGTTCTCCC	1695	2716	CTTTGGATTTGTGCTTAAGAAATTAATGCAAAAGAACCAAGACAGATGATATGAGA	2775
Db	1649	ATGACCCCAAGTCTGCTTATGCTGCTCACTAGTACATTTTCAAGGGCTCCCTGTTCCC	1708	2727	GCITTTGATCATGCC--AAGCCTGAAAGTGTCAAAGAGCAGATACAGAAATGTCATGAGA	2784
Qy	1696	TCCTTCATGACAGAGAGGATCTTGTATTTGAGTCTAAATTAATTTATTTGAGTAGATG	1755	2776	ATTGCTTTTAAATTTGTAAATTTAAAAAATTTAGATCGTTACTTTGGAATTTGAGCCCTA	2835
Db	1709	TGCTTCATGAACAGAGAGAAATTTCTGACTTGCAGTCTAAATTAATTTATTTGCGGTAG	1768	2785	ATTGCTTTTAAATTTGTAAATTTAAAAAGCCCTTAATTTGCTGGGCTGGGTTCAA--TC	2840
Qy	1756	TTGCCAAAGGATGAGTACCTTCACACCTGACACAGCCAAATTTATACATCGTACTTGA	1815	2836	AGTCTGTGGTGGACAGATAAATTAATTTTCTGGGTGAAATTAATTTAGACTTTGTTT	2895
Db	1769	TTGCCAAAGGATGAGTACCTTCACACCTGACACAGCCAAATTTATACACCGGACCTGA	1828	2841	TGTGTAGATAGTGGGTTGACCCCTTATGATTTTGTAGACCAAACTGTGTGGGCTTGTGT	2900
Qy	1816	ACAGTCACAATATTTCTTCTATGAGGATGGGATGCTGTGGTGGCAGATTTTGGAGAT	1875	2896	TGACAG-CTATGGGTTTATTTCTTAGAACATTTCTTCTTCTTCTTCTTCTTCTTCTTCT	2954
Db	1829	ACAGCCACAATATTTCTGCTCTATGAGGATGGCCATGCTGTGGTGGCAGATTTTGGAGAT	1888	2901	TGAGGGTCTCCTGTGGGTTTCTTAAAAACAGCTGGCTGATTTATCTCTCTTCTTCTT	2960
Qy	1876	CAAGATTTCTCAGTCTGAGTGAAGACATGACAAACACCTTGGGAACTCCGCTT	1935	2955	TCTAGTGTTCACCTCTGTGATTAAGATTTCTTTGGTGAATTAAGAAAAAATAAAAAA	3014
Db	1889	CAAGATTTCTGAGTCCCTGAGTGAAGACATGACAAAGCAGCCAGGAACTGGCT	1948	2961	----GTTTGTACTTCTGTGATTAAGTCTCTTGGTGTCTAGAAAAAATAAAAAA	3016
Qy	1936	GGATGGCTCCTGAGTGTTCAGCGATGCTACCTCGGTACACCATCAAGCAGATGTCTTCA	1995	3015	AGGGCGGC 3022	
				3017	GGGGCGGC 3024	

## RESULT 7

US-09-947-199-9  
; Sequence 9, Application US/09947199  
; Patent No. US20020127684A1  
; GENERAL INFORMATION:  
; APPLICANT: Raju, Jeyaseelan  
; TITLE OF INVENTION: NOVEL CARP PROTEIN AND NUCLEIC ACID MOLECULES AND USES  
; FILE REFERENCE: THEREFOR  
; FILE REFERENCE: MN-068CP2  
; CURRENT FILING DATE: 2001-09-05  
; PRIORITY FILING DATE: 2001-09-05  
; PRIORITY FILING DATE: 1998-12-11  
; PRIORITY FILING DATE: 1999-04-14  
; PRIORITY FILING DATE: 1999-04-14  
; PRIORITY FILING DATE: 1999-04-14  
; PRIORITY FILING DATE: 1999-12-10  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 9  
; LENGTH: 2505  
; TYPE: DNA  
; ORGANISM: Rattus norvegicus  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(2505)  
US-09-947-199-9

Query Match 62.6%; Score 1892.2; DB 9; Length 2505;  
Best Local Similarity 84.7%; Pred. No. 0;  
Matches 2122; Conservative 0; Mismatches 383; Indels 0; Gaps 0;  
  
QY 48 ATGGGAAATTAATCTAGACCAACCCAACTTGTACTGATGAATGGAAGAAAAGTC 107  
DB 1 ATGGGAAATTAATCTAGACCAACCCAACTTGTACTGATGAATGGAAGAAAAGTC 60  
  
QY 108 AGTGAATCATATGTATTATCAATAGAAAGATAGAAAGATGACCTCAGATCAAGGAAAA 167  
DB 61 AGTGAATCATATGTATTATCAATAGAAAGATAGAAAGATGACCTCAGATCAAGGAAAA 120  
  
QY 168 GAATCAGAGAACTAAGGAATATATTTGGCTCTGATGAAGCCCTCAGTAAAGTCAATTTA 227  
DB 121 GAATTTCAAGAACTAAGGACATCTTTGGCTCTGATGAAGCCCTCAGTAAAGTCAATTTA 180  
  
QY 228 AATTACGACATGAATAGGCTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 287  
DB 181 AATTACGACATGAATAGGCTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 240  
  
QY 288 AATCATATATTCGAATCTTATTTGAAAGGCTCCGCCATCTCGACTGACAGAAAT 347  
DB 241 AAGTCATATATTCGATCTTATTTGAAAGGCTCCGCCATCTCGACTGACAGAAAT 300  
  
QY 348 GGATTTACAGCTTGCATTTAGCAGTTTACAGGATATGAGATTTGATTTGATTTGATTTGATTTG 407  
DB 301 GGATTTTCCAGCTTGCATTTAGCAGTTTACAGGATATGAGATTTGATTTGATTTGATTTGATTTG 360  
  
QY 408 CTTTACAGTGGAGCTGATATACAGAGGTTGATACGGTGGCTCAGTCCCTCCATATT 467  
DB 361 TTGCACAGGGAGCAGATGTTGAGAGGTTGATACGGTGGCTCAGTCCCTCCATATT 420  
  
QY 468 GCTACAAATAGTGGCCACCTAGAGGCTGTGATGTGTTGCAACATGAGGCTTAATGTC 527  
DB 421 GCTGCAATAGTGGACACCCAGAGGCTGAGAGGCTGCTGCTACAAATGGGGCCCAATGTG 480  
  
QY 528 AATATTCAAGATGCAATTTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 587  
DB 481 AATTTTCAAGATGCAATTTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 540  
  
QY 588 CAGGTAACCTGGCTTCTTTTGAATTTGGTGTGATGTAATGTAAGTGAAGTTTGA 647  
DB 541 CAGGTAACCAAGTGTCTTTTGAATTTGGTGTGATGTAATGTAAGTGAAGTTTGA 600

QY 648 GATAGACCCCTCCACCTAGCATCTCAGAAAGGATTTTGAATATTGCAAACTCTTGATG 707  
DB 601 GACAGGCTCTGCACTGCGCTCTGCAAGGCTTCTTCAACATTTGAAACTCTCTGTA 660  
  
QY 708 GAAGAGGAGCAGCAAGACAGATGTAATGCTCAAGATATGAGCAATGAGCAATGCTCCACTCAT 767  
DB 661 GAAGAGGAGCAGCAAGACAGATGTAATGCTCAAGATATGAGCAATGAGCAATGCTCCACTCAT 720  
  
QY 768 TTCTGTTCTCGATTTGGACACCATGATATAGTAAAGTATCTGCTGCAAGATGATTTGAA 827  
DB 721 TTCTGTTCTCGATTTGGACACCATGATATAGTAAAGTATCTGCTGCAAGATGATTTGAA 780  
  
QY 828 GTTCAACCTCATGTTGTTAATATCTATGAGATACCCCTTACCTGGCATGCTACAT 887  
DB 781 GTCCAGCCTCAGCTCATTAACATCTATGTTGACATCTCTTTGACCTGGCATGCTACAT 840  
  
QY 888 GCAAAATTTGAAAGTTGCCAAGGAAATCATCAAAATATCAGGAACAGAAAGTCTGACTAAG 947  
DB 841 GCAAAATTTGAAAGTTGCCAAGGAAATGTCCAGGTAACAGGAATCTGAAAGTCTGACTAAG 900  
  
QY 948 GAAACATCTTCAGTGAACAGCTTTTTCATAGTGTCTTATCTATGGAAGCAAGATTTGAC 1007  
DB 901 GAAACATCTTCAGTGAACAGCTTTTTCATAGTGTCTTATCTATGGAAGCAAGATTTGAC 960  
  
QY 1008 CTAGTCARATTTCTTCTTCTGATCAGATGTCATAACATCAACCAAGGAGGATGGG 1067  
DB 961 CTGTCARATTTCTTCTTCTGATCAGATGTCATAACATCAACCAAGGAGGATGGG 1020  
  
QY 1068 CACATGGATTAACATCTCTGTTGTTACCAAGGTCATCTGCTGCTGCTGCTGCTGCTGCTGCTG 1127  
DB 1021 CACATGGATTAACATCTCTGTTGTTACCAAGGTCATCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080  
  
QY 1128 GATAATGGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1187  
DB 1081 GATAATGGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1140  
  
QY 1188 GATGACACACATGTTTGTATGTTGGCTTTATGAAAGGCGCATGATGCTATGTCACATCTC 1247  
DB 1141 GATGACACACATGTTTGTATGTTGGCTTTACGAGAAAGGACATGATGCTATGTTTACATCTC 1200  
  
QY 1248 CTGAGCATTATTAAGAGACCAAGATGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1307  
DB 1201 CTGAGCATTATTAAGAGACCAAGATGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1260  
  
QY 1308 GGAGATGGCTCTCTATGCTGCTTCCATCACCCTGGGAAAGATTAAAGCATGACAAAA 1367  
DB 1261 GGAGATGGCTCTCTATGCTGCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTT 1320  
  
QY 1368 GAGAGGAGATATTTCTCTCTAAGAGCTGATGCTTCCATTTCCATTTCCATTTCCATTTCCATTT 1427  
DB 1321 GAGAGGAGATATTTCTCTCTAAGAGCTGATGCTTCCATTTCCATTTCCATTTCCATTTCCATTT 1380  
  
QY 1428 TCAGAAATTTGAGTTCATGAGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1487  
DB 1381 TCAGAAATTTGAGTTCATGAGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1440  
  
QY 1488 CGATGCAAAATTAATAGTGGCTATAAAAGCTTATCGAGCAATACCTACTGCTCCAAAG 1547  
DB 1441 CGATGCAAAATTAATAGTGGCTATAAAAGCTTATCGAGCAATACCTACTGCTCCAAAG 1500  
  
QY 1548 TCAGATGAGATATGTTTGGGAGAGGTTGCTTCCATTTCTCTGCGAGCTCAATCATCTCTGCTG 1607  
DB 1501 TCAGATGAGATATGTTTGGGAGAGGTTGCTTCCATTTCTCTGCGAGCTCAATCATCTCTGCTG 1560  
  
QY 1608 GTAATTTGAGTGGCTGCTTGTGATGATCCAGCCAGTTTGGCATTTGCTGCTGCTGCTGCTGCTG 1667  
DB 1561 GTAATTTGAGTGGCTGCTTGTGATGATCCAGCCAGTTTGGCATTTGCTGCTGCTGCTGCTGCTG 1620  
  
QY 1668 TACATATCAGGGGTTCTCTGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1727  
DB 1621 TACATATCAGGGGTTCTCTGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1680

QY 1728 CAGTCTAAATTAATTAATTCAGTAGATGTTGCCAAGGCATGGAGTACCTTCACAACTG 1787  
DB 1681 CAGTCTAAATTAATTAATTCAGTAGATGTTGCCAAGGCATGGAGTACCTTCACAACTG 1740  
QY 1788 ACAGAGCAATATACATCGTACTTGAACAGTCACAAATATTTCTCTATGAGGATGG 1847  
DB 1741 ACCAGCAATATACATCGTACTTGAACAGTCACAAATATTTCTCTATGAGGATGG 1800  
QY 1848 CATGCTGTGGTGGAGATTTTGGAGATCAAGATTTCTACAGTCTCTGGATGAAGACAAC 1907  
DB 1801 CATGCTGTGGTGGAGATTTTGGAGATCAAGATTTCTACAGTCTCTGGATGAAGACAAC 1860  
QY 1908 ATGACAAACAACTGGGAACCTCCGTTGGATGGCTCTCGAGTGTTCAGCAGTGCAT 1967  
DB 1861 ATGACAAACAACTGGGAACCTCCGTTGGATGGCTCTCGAGTGTTCAGCAGTGCAG 1920  
QY 1968 CGGTACACATCAAGCAGATGCTTTCAGCTATGCTCTGTGTGTGGGAAATTCCTACT 2027  
DB 1921 AGATACACATCAAGCAGTGTCTTTCAGTACTCCCTGTGTGTGGGAGTCTCTACT 1980  
QY 2028 GCGAAATTCATTCGCTCACTCAAGCAGCAGCTGCGCAGCAGATGGCTTACCAAC 2087  
DB 1981 GGAGAAATTCATTCGCTCACTCAAGCAGCAGCTGCGCAGCAGATGGCTTACCAAC 2040  
QY 2088 CACATCAGACCTCCATTCGCTATTCATTCAGCAGCAGCTGCGCAGCAGATGGCTTACCAAC 2147  
DB 2041 CACATCAGACCTCCATTCGCTATTCATTCAGCAGCAGCTGCGCAGCAGATGGCTTACCAAC 2100  
QY 2148 GGGTGGAAATCGATGCTTGAAGGAGACCCGAATTTTCTGAAGTGTTCATGAAGTTAGAA 2207  
DB 2101 GGGTGGAAATCGATGCTTGAAGGAGACCCGAATTTTCTGAAGTGTTCATGAAGTTAGAA 2160  
QY 2208 GAGTGTCTCTGCAATTCAGTGTCTTCCGATCAATGATCAAGCAGTGTTCATGAAGTTAGAA 2267  
DB 2161 GAGTGTCTCTGCAATTCAGTGTCTTCCGATCAATGATCAAGCAGTGTTCATGAAGTTAGAA 2220  
QY 2268 TCACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2327  
DB 2221 TCACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2280  
QY 2328 GCAGCATTAAGAGTCTGTTTGAATTCGAATATGCTCTTAAATGCAAGTCTCTATGCTGT 2387  
DB 2281 GCAGCATTAAGAGTCTGTTTGAATTCGAATATGCTCTTAAATGCAAGTCTCTATGCTGT 2340  
QY 2388 TTGTCCTCAAGTCTGAGCATATTCCTCTCAAGTCTCTCTTGGAGGAGATGAAGA 2447  
DB 2341 TTGTCCTCAAGTCTGAGCATATTCCTCTCAAGTCTCTCTTGGAGGAGATGAAGA 2400  
QY 2448 AGTCTTCAATACACACCAATTCACAAATATGCTATGCTATGCTATGCTATGCTATGCT 2507  
DB 2401 AGTCTTCAATACACACCAATTCACAAATATGCTATGCTATGCTATGCTATGCTATGCT 2460  
QY 2508 CATTTTCAATTCCTGCGAAATAGTAGAGCTTTTGGAGGAGATGAAGA 2552  
DB 2461 CATTTTCAATTCCTGCGAAATAGTAGAGCTTTTGGAGGAGATGAAGA 2505

## RESULT 8

US-10-626-173-9  
; Sequence 9, Application US/10626173  
; Publication No. US20040110232A1  
; GENERAL INFORMATION:  
; APPLICANT: Raju, Jeyaseelan  
; TITLE OF INVENTION: NOVEL CARB PROTEIN AND NUCLEIC ACID MOLECULES AND USES  
; FILE OF INVENTION: THEREFOR  
; FILE REFERENCE: MNI-068CP2  
; CURRENT APPLICATION NUMBER: US/10/626.173  
; CURRENT FILING DATE: 2003-07-24  
; PRIOR APPLICATION NUMBER: US/09/947,199A  
; PRIOR APPLICATION NUMBER: US 60/111,936  
; PRIOR FILING DATE: 1998-12-11  
; PRIOR APPLICATION NUMBER: US 09/291,839  
; PRIOR FILING DATE: 1999-04-14

; PRIOR APPLICATION NUMBER: US 09/458,457  
; PRIOR FILING DATE: 1999-12-10  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 9  
; LENGTH: 2505  
; TYPE: DNA  
; ORGANISM: Rattus norvegicus  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(2505)  
US-10-626-173-9

Query Match 62.6%; Score 1892.2; DB 17; Length 2505;  
Best Local Similarity 84.7%; Pred. No. 0;  
Matches 2122; Conservative 0; Mismatches 383; Indels 0; Gaps 0;  
QY 48 ATGGGAATTAATACTAGACCAACCCAAACTTGTACTGTATGAATGGAAGAAAAATGC 107  
DB 1 ATGGGAATTAATACTAGACCAACCCAAACTTGTACTGTATGAATGGAAGAAAAATGC 60  
QY 108 AGTAATCATATGTTATCATATAGAAAGATTAGAAAGATGACCTGACATCAAGAAAAA 167  
DB 61 AGTAATCATATGTTATCATATAGAAAGATTAGAAAGATGACCTGACATCAAGAAAAA 120  
QY 168 GAATGACAGAACTAAGGAATATATTTGCTCTGTATGAAGCCCTTCAGTAAAGTCAATTTA 227  
DB 121 GAATTCAGAACTAAGGACATCTTTGCTCTGTATGAAGCCCTTCAGTAAAGTCAATTTA 180  
QY 228 AATTACCGACTGAAATGGCTGTCTTACTTCAATTTAATGTTGATTTGAGGAGCAAG 287  
DB 181 AATTACCGACTGAGGCTGTCCCTGTCTACACCTCTGTGTGTCTGTGGCGCAAC 240  
QY 288 AATACATATTCGAATCTTATGTTGAAGGCTCCGCCCTCTCGACTGACAGAAAT 347  
DB 241 AAGTCAATATCCGTGCTTATGTTAAAGGCTCCGCCCTCTCGACTGACAGAAAT 300  
QY 348 GGATTTACAGCTTGTGATTTAGCAGTTTACAAAGGATTAATGCAAAATGATCACTTCTCTG 407  
DB 301 GGGTTTCCACTCTGCACCTGGCCGTTTACAAAGGACAGCCCGAACTTATCACTTCACTG 360  
QY 408 CTTACAGTGGAGCTGATATACAGAGTTGATAGAGTGGCTCACTGCCCTCCATATT 467  
DB 361 TTGACAGCGGAGCAGATGTTTACAGAGTGGATAGGCTTGGCTTACAGCCCTCCACATA 420  
QY 468 GCTACATAGCTGGCACTAGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 527  
DB 421 GCTGCAATAGCTGGACACCCAGAGGCTGCAAGGCTGCTACAAATGGGSCCAATGTG 480  
QY 528 AATATCAAGATGAGTTTTTTTCACTCCATTTGATATTTGAGCGTACTATGGACATGAA 587  
DB 481 AATGTTCAAGATGCGCTTCTTTCACCCCACTGCAATTTGAGCGCTTCTATGGGACGAG 540  
QY 588 CAGTAACTCGCTTCTTTTGAATTTGTTGCTGTATGTAATGTAAGTGGTCAAGTTGGA 647  
DB 541 CAGTAACTCGCTTCTTTTGAATTTGTTGCTGTATGTAATGTAAGTGGTCAAGTTGGA 600  
QY 648 GATAGACCCCTCCACCTAGCATCTGCAAAAGGATTTTGAATTTGCAAAATCTTTGATG 707  
DB 601 GACAGGCTCTGCACTGGCTCTGCAAAAGGCTTCTTCAACATTTGTGAACTCTCTGTA 660  
QY 708 GAAGAGGAGCAAGCAGATGTAATGCTCAAGATAATGAAGACATGTCCTCCACTCCAT 767  
DB 661 GAAGAGGAGCAAGCAGATGTAATGCTCAAGATAATGAAGACATGTCCTCCACTCCAT 720  
QY 768 TTCTGTTCTCGATTGGACACCAATGATAGTAAATGATCTGCTCTCAAAAGTGAATGGAA 827  
DB 721 TTCTGTTCTCGATTGGACACCAATGATAGTAAATGATCTGCTCTCAAAAGTGAATGGAA 780  
QY 828 GTTCAACCTCATGTTGTTAATATCTATGAGATACCCCTTACCTGATGCTGCTACAT 887  
DB 781 GTTCAACCTCATGTTGTTAATATCTATGAGATACCCCTTACCTGATGCTGCTACAT 840

QY 888 GGCAAAATTTGAAGTTCGCAAGGAATCATCCAAATATACAGGAACAGAAAGTCTGACTAAG 947  
Db 841 GGAATTTTGAAGTTCGCAAGGAATTTGTCAGAGTAAACAGAACTCAAGTCTGACTAAG 900  
QY 948 GAAAAACATCTTTCAGTGAACAGAGCTTTTTCATAGTCTTTGACATGTTGACATGGAAGAGCAATTCAC 1007  
Db 901 GAAAAACATCTTTCAGAGAGAGAGCTTTTTCATAGTCTTTGACATGTTGACATGGAAGAGCAATTCAC 960  
QY 1008 CTAGTCAAAATTTCTTTGATCAGAATGTCATAAACATCAACCAACAGAAAGAGGATGGG 1067  
Db 961 CTGCTCAAAATTTCTTTGATCAGAATGTCGTGAACATTAACCAACAGAAAGAGATGGG 1020  
QY 1068 CACATGGGATTAACATCTGCTTCTACACAGGTCACATTCGCTGGTTCAGTCTTCTTACTG 1127  
Db 1021 CACAGGATTTGACATCTGCTTCTACACAGGCTATACCGCTGGTTCAGTCTTCTTACTT 1080  
QY 1128 GATAATGGAGCTGATATGAATCTAGTGGCTTTGTGATCCACAGAGTCTAGTGGTGAAGAAA 1187  
Db 1081 GATAATGGTGCAGATATGAATCTTGTGGCTTTGTGATCCACAGAGTCTAGTGGTGAAGAAA 1140  
QY 1188 GATGCGCAGACATGTTGATGCGGCTTATGAAAAGGCGCATGATGCCATTTGTCACATC 1247  
Db 1141 GATGAGCAGACATGTTGATGCGGCTTACGAGAAGGACATGATGCCATTTTACATC 1200  
QY 1248 CTGAAGCAATTAAGAGACACAGATGAATTTGCCCTGTAATGAATATTTCTCAGGCTGGA 1307  
Db 1201 CTGAAGCACTACAGAGACCCAGGAGGAGCTGCAATGTAACGAATATTTCCAGGCTGGA 1260  
QY 1308 GAGAGATGGCTCCTATGTCCTTCAAGAGTGGATGCTTCAATTTCCATTTCCATCTCAGCTC 1367  
Db 1261 GAGAGATGGCTCCTATGTCCTTCAAGAGTGGATGCTTCAATTTCCATTTCCATCTCAGCTC 1320  
QY 1368 GAGAGGACAGATATTTCTCTTCAAGAGTGGATGCTTCAATTTCCATTTCCATCTCAGCTC 1427  
Db 1321 GAGAGGACAGATATTTCTCTTCAAGAGTGGATGCTTCAATTTCCATTTCCATCTCAGCTC 1380  
QY 1428 TCAGAAATTCAGTTCATGATGATTTATGGCTCAGGTTCTTTTGGGAAAGTATATAAGGA 1487  
Db 1381 TCCGAAATCGAGTTCACAGAGATATTCGGCTCGGTTCTTTTGGGAAAGTATATAAGGG 1440  
QY 1488 CGATGCAGAAATTAATAGTGGCTATATAACGTTATCGAGCAATACCTACTGCTCCAG 1547  
Db 1441 CGATGCAGAAATTAATAGTGGGATCAACAGATCCGAGCAACACCTACTGCTCCAG 1500  
QY 1548 TCAGATGATGATATTTTCCGAGAGGTGCTCAATTTCTGCGAGCTCAATCATCCCTGC 1607  
Db 1501 TCAGAGTGTGATATTTTCCGAGAGGTGCTCAATTTCTGCGAGCTCAACCAACCTCTC 1560  
QY 1608 GTAATTCAGTTTGTGGGTGCTTCTGTAATGATCCAGCCAGTTTGCATTTGTCTACTCAA 1667  
Db 1561 GTGGTTTCAGTTTGTGGGTGCTTCTGTAATGATCCAGCCAGTTTGCATTTGTCTACTCAG 1620  
QY 1668 TACATATCAGGGGTTCTGTTTCTCCCTTTCATGACAGAGAGGATTTCTTGAATTTG 1727  
Db 1621 TACATTTACAGAGGCTCCCTGTTTCTCCCTGCTTCATGAACAGAGAAATTTCTGACTTG 1680  
QY 1728 CAGTCTAAATTAATTTAGTGTAGATGTTGCAAAAGGATGAGTACCTTCAACACCTG 1787  
Db 1681 CAGTCTAAATTAATTTAGTGTGTAGATGTTGCAAAAGGATGAGTACCTTCAACAGCTTG 1740  
QY 1788 ACACAGCCAAATTAACATCGTGTACTTGAACAGTCAATATTTCTTCTATGAGATGGG 1847  
Db 1741 ACCAGCCAAATTAACATCCGCGACCTGGAACAGCCCAATATTTCTGCTATGAGGATGGC 1800  
QY 1848 CATGCTGTGGTGGCAGATTTTGGAGATCAAGATTTTCAAGTCTCTGATGGAAGACAAC 1907  
Db 1801 CATGCTGTGGTGGCAGATTTTGGAGATCAAGATTTTCTGAGTCCCTGGATGAAGACAAC 1860  
QY 1908 ATGACAAAAACAACCTGGGAACCTCGTTGGATGGCTCCTGAGGTGTTTACGAGTGTCACT 1967  
Db 1861 ATGACAAAGCAGCAGGGAACCTGCGTGGATGGCCCTGAGGTGTTTACACAGTGTCAAG 1920  
QY 1968 CGGTACACCATCAAGCAGATGCTTTCAGCTATGCTCTGCTGTGGGAAATTTCTCACT 2027

Db 1921 AGATACACCATCAAGGCTGATGCTTTCACTTACTCCCTGTGTCTGGAGCTCTCTACT 1980  
QY 2028 GGGAAATTTCCATTTCTGCTCATCTCAAGCCAGCGGTGCGGACAGACATGGCTTACCAC 2087  
Db 1981 GGAATTTTCCATTTCTGCTCATCTCAAGCCAGCGGTGCGGACAGACATGGCGTATCAC 2040  
QY 2088 CACATCAGACCTTCCATTTGGCTTATTCATTTCCCAAGCCCATATCATCTCTCTGTATAGA 2147  
Db 2041 CACATCAGACCCGCCATCGGCTTATTCATTTCCCAAGCCCATCTCATCTCTGTATCGG 2100  
QY 2148 GGTGGAACGCGATGCTCTCAAGGAAGACCCGAATTTTCTGAAGTTGTCTATGAAGTTAGAA 2207  
Db 2101 GGTGGAATGCGATGCTCTCAAGGAAGACCCAGATTTCTCTGAAGTCTGTAGCAAACTGGAG 2160  
QY 2208 GATGTCCTCTGCAACATTTAGCTGATGTCCTCTGATCAAGTAAACAGAGTGGTCTCTC 2267  
Db 2161 GAGTGCCTATGCAATGTGAGCTCATGTCCTCAGCATCAAGTAAACAGAGTGGCTCTCTG 2220  
QY 2268 TCACCTTCTTCTTCTTCTGATTTGCCCTGTTGAACCCGGGAGGACCTTGGCGGAGTCACTG 2327  
Db 2221 TCACCTTCTTCTTCTTCTGATTTGCCCTGTTGAACCCGGGAGGACCTTGGCGGAGTCACTG 2280  
QY 2328 GCAGCATTAAGAGTCTGTTTCGAATTTGGAATATGCTCTAAATGCAAGTCTCTATGCTGCT 2387  
Db 2281 GCAGCCTTACGGAGCGGTTTTCAGTTGAGTATGCCCTAAATGCAAGTCTCTATGCTGCTG 2340  
QY 2388 TTGTCCTCAAGTCTGAGCAATATTTCTCTCAAGTCTGCTTTTGGAGGAGATGAAAGA 2447  
Db 2341 TGTGTCCTCAAGTCTGAGCAACACTCTTAATCCGGGCTGCTTTTGGAGGAGATGAAAG 2400  
QY 2448 AGTCTTCAATACACACCCATTTGACAAATATGCTATGATCCGATCCCATGAGTCAATG 2507  
Db 2401 AGCACCCAGTATTTCAACTCTTGACAAATACGGCTATGTCGTCTGATCCCATGAGCCTGACG 2460  
QY 2508 CATTTTCATTTCTGCGAAATAGTAGCAGCTTTTGAGGACAGCAGC 2552  
Db 2461 CACCTTCACTCCCGCAAGACGACAGCAACTTTTGAGGACAGCAAC 2505

## RESULT 9

US-09-833-381-1151  
; Sequence 1151, Application US/09833381  
; Patent No. US20020132090A1  
; GENERAL INFORMATION:  
; APPLICANT: Robison, Keith E.  
; TITLE OF INVENTION: No. US20020132090A1el Nucleic Acid and Protein Homologs  
; FILE REFERENCE: 5900-119  
; CURRENT APPLICATION NUMBER: US/09/833,381  
; CURRENT FILING DATE: 2001-04-11  
; PRIOR APPLICATION NUMBER: 09/516,448  
; PRIOR FILING DATE: 2000-02-29  
; NUMBER OF SEQ ID NOS: 2050  
; SOFTWARE: Fast-Seq for Windows Version 3.0  
; SEQ ID NO 1151  
; LENGTH: 740  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)..(740)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-833-381-1151

Query Match 14.1%; Score 427.6; DB 9; Length 740;  
Best Local Similarity 78.4%; Pred. No. 1.7e-106;  
Matches 519; Conservative 0; Mismatches 142; Indels 1; Gaps 1;

QY 16 CGGCGCTGGAGAGAGAAAGAACTTATAATGGAATTTATAATCTAGACCAACCC 75  
Db 41 CAGCACAGGAGAGAAAGAAAGACTTCTTTAAATGGGAATTTACAAATCCAGCACAC 100  
QY 76 AAACCTGTACTGATGAATGGAAGAAAAAGTCAGTGAATCATATGTTTATCACAATAGAAA 135

```
Db 101 AGACTTGTCTGATGAATGGAAGAAAGTTAGTGAATCTTACGCTATTATCATAGAAA 160
QY 136 GATTAGAAGATGACCTGCAGATCAAGGAAAGAAAGTACAGAACTAAGGAAATATTG 195
Db 161 GGCTGAGGATAACCTGCAGATCAAGAAATGAATTTCAAGAACTAAGSCACATCTTGG 220
QY 196 GCTCTGATGAAGCCTTCAGTAAAGTCAATTTAAATTAACGACCTGAAAATGGGCTGTTC 255
Db 221 GCTCTGATGAAGCCTTCAGTAAAGTCAATTTAAATTAACGACCTGAAAATGGGCTGTTC 280
QY 256 TACTTCAATTTATGTTGCAATTTGGAGGCAAGAAATACATATTCGAATCTTATGTTGA 315
Db 281 TCGTACACCTCTGCTGTCTGTGGCGGCAACAAGTCACATATCCGTGCGCCCTTATGTTAA 340
QY 316 AAGGCGCTCGGCCATCTCGACTGACAGAAATGGAATTTACAGCCTTGCAATTTAGCAGTTT 375
Db 341 AAGGCGCTCGGCCATCTCGACTGACAGAAATGGAATTTACAGCCTTGCAATTTAGCAGTTT 400
QY 376 ACAAGGATAATGCAGAAATGATCACTTCTCTGCTTAC-AGTGGAGCTGATATACAGCAG 434
Db 401 ACAAGGACAGCCCGGAATTTATCACTTCACTGTTGGCAAGCGGAGCAGATGTTACGCAA 460
QY 435 GTTGATACGGTGGCCTCACCTGCCCTCATATTTGCTCAATAGTCTGACCACTAGAGGCT 494
Db 461 GTGGATACGGTGGCCTCACAGCCTCCACATAGCTGCAATAGCTGACACCCAGAGGCT 520
QY 495 GCTGATGCTGTGTTGCAACATGGAGCTAATGTCATTAATTTCAAGATGAGTTTTCAC 554
Db 521 GCANAAGTGTCTGTTGCAACATGGGCGCAACGTTGAATGTTCAAGATGCGCTTCTTTCACC 580
QY 555 CCATTGCAATTTGAGCTGCTATGAGCATGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 614
Db 581 CCACTGCAATTTGAGCTGCTATGAGCATGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 640
QY 615 GGTGCTGATGTAATGTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 674
Db 641 GGGGTGATGTCNATGTAAGCGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 700
QY 675 AA 676
Db 701 AA 702

RESULT 10
US-09-833-381-1149
; Sequence 1149, Application US/09833381
; Patent No. US20020132090A1
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. US20020132090A1el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1149
; LENGTH: 616
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-833-381-1149
Query Match 14.1%; Score 427.4; DB 9; Length 616;
Best Local Similarity 81.8%; Pred. No. 1.7e-106;
Matches 505; Conservative 0; Mismatches 111; Indels 1; Gaps 1;
```

```
QY 338 GACAGAAATGGAATTTACAGCTTGCATTTAGCAGTTTACAGGATAATGCAGAAATGAT 397
Db 1 GACGAGAAATGGTTTCCAGCTTGCACCTGGCGGTTTACAGGACAGCCGGAATTTAT 60
QY 398 CACTTCTCTGCTTACAGTGGAGTGTATATACAGCAGGTTGGATAGCGTGGCTCACTGC 457
```

```
Db 61 CACTTCACTGTGTACAGCGGAGCAGATGTTTACAGAAAGTGGGATACGGTGGCTCACAGC 120
QY 458 CCTCCATATTGCTACAAATAGCTGGCCACCTAGAGGCTGTGTATGTCTGTTTGCACATGG 517
Db 121 CCTCC-CATAGCTGCAATAGCTGACACCCAGAGGCTGCAGAAAGTGTCTGCTSCAACATGG 179
QY 518 AGCTAATGTCATATTCAAGATGCAAGTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 577
Db 180 GGCCACAGTGAATGTTCAAGATGCGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 239
QY 578 TGGACATGAACAGGTAACCTGCGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 637
Db 240 TGGCAGCAGCAGGTAACCAAGTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 299
QY 638 TGAAGTTGGAGTAGAGCCCTCCACTAGCATCTGCAAAAGGATTTCTTGAATATTGCAAA 697
Db 300 TGAAGTTGGGACAGGCTCTGCACCTGCGCTCTGCAAAAGGCTTCTTCAACATTTGTGAA 359
QY 598 ACTTCTGATGAAGAGGAGCAGCAAGCAGATGTGATGCTCAAGATATGAAGAGCAATGT 757
Db 360 ACTCCTGTAAGAGAGGAGGAGCAAGCAGATGTGAACGCTCAGACATGAAGACCACTGT 419
QY 758 CCCACTCCATTTCTGTCTGATTTGGACACCATGATATAGTTAAAGTATCTGCTCCAAAG 817
Db 420 CCCTCTGCACTTCTGTCTGATTTGGACACCAACATATAGTGAGCTACCTGCTCCAGAG 479
QY 818 TGATTTGGAAGTTCAACCTCATGTTTAAATATCTATGAGATACCCCTTACACCTGGC 877
Db 480 TGACTTAGAGGTCGAGCCTCACGTCAATTAACATCTATGTGTGACATCTCTTTCACCTGGC 539
QY 878 ATGTCTACAATGGCAATTTGAAGTTCGCAAGGAAATCATCCAAATATCAGGAACAGAAAG 937
Db 540 ATGTCTACAATGGAAATTTGAAGTTCGCAAGGAAATTTCCAGGTAACAGGAACTGAAG 599
QY 938 TCTGACTAAGGAAACCA 954
Db 600 TCTGACTAAGGAAACCA 616

RESULT 11
US-09-833-381-1150
; Sequence 1150, Application US/09833381
; Patent No. US20020132090A1
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. US20020132090A1el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1150
; LENGTH: 736
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; LOCATION: (1)-(736)
; OTHER INFORMATION: n = A,T,C or G
US-09-833-381-1150
Query Match 13.7%; Score 414.6; DB 9; Length 736;
Best Local Similarity 81.2%; Pred. No. 6.6e-103;
Matches 480; Conservative 0; Mismatches 111; Indels 0; Gaps 0;
```

```
QY 312 TTGAAGGGCTCGCCCATCTCGACTGACAGAAATGGAATTTACAGCCTTGCATTTAGCA 371
Db 1 TTAAGAGGGCTCGCTCCATCCAGACTGACAGAAATGGTTTCCAGCTCTGACCTGGCC 60
QY 372 GTTTACAAGGATAATCAGAAATGATCACTTCTCTCTTTCACAGTGGAGCTGATATACAG 431
```

Db 61 GTTTAAGGACGACCGGAACTTATCACTTCTACTGTTCACAGCGGACGAGATGTTACG 120  
Qy 432 CAGGTTGGATACGGTGGCCCTCACTGCGCTCCATATTTGCTACATAGCTGGCCACCTAGAG 491  
Db 121 CAAGTGGGATACGGTGGCCCTCAAGCCCTCCATAGCTGCAATAGCTGGACACCCAGAG 180  
Qy 492 GCTGCTCATGTGTTGTTGCAACATGAGGCTAATGTCAATATTCAAGATGCAAGTCTTTTTC 551  
Db 181 GCTGCAAGATGCTGTACAAATGGGGCCATGTGATTTCAAGATGCCGCTTCTTC 240  
Qy 552 ACTCCATTGCAATATGAGCGGTACTATGGACATGAACAGGTAACTCGCCTTTTGTGAAA 611  
Db 241 ACCCCACTGCACATGTCAGGCTACTATGGCCACGACGAGTAACCAAGTCTCCTTTTGAAG 300  
Qy 612 TTTGGTCTGATGTAATGTAAGTGGTGAAGTTGGAGTAGAGCCCTCCACCTACATCT 671  
Db 301 TTTGGTCTGATGTAAGTGGTGAAGTTGGGACAGGCTCTGCACCTGGCCTCT 360  
Qy 672 GCAAAAGGATCTTTGATATTCGAAACTCTTGTATGGAAGAGGACGCAAGCAGATGTG 731  
Db 361 GCAAAAGGCTTCTTCAACATTTGAACTCTCTGTAGAAAGGAGGACAAACGAGATGTG 420  
Qy 732 AATGCTCAAGATATGAAGACCATGTCCCACTCCATTTCTGTTCTCGATTTGGACACCAT 791  
Db 421 AAGCTCAGGACAAATGAAGACCAAGCTCCCTCTGCACCTTCTGATTTGGACACCAAC 480  
Qy 792 GATATAGTTAAGTATCTGCTGCAAGTGAATTTGGAAGTTCAACCTCATGTTTAAATATC 851  
Db 481 AATATAGTGAAGTACCTGCTCCAGATGACTTANAGTCCAGCTCACGTCAATTAACATC 540  
Qy 852 TATGGAGATACCCCTTACACCTGGGATGTACAAATGGGCAATTTGAAGTT 902  
Db 541 TATGGTGACATCTCTTTGCACTGGCATGTACAATGGAAATTTTGAANTT 591

## RESULT 12

US-09-833-381-1147/c  
; Sequence 1147, Application US/09833381  
; Patent No. US20020132090A1  
; GENERAL INFORMATION:  
; APPLICANT: Robison, Keith E.  
; TITLE OF INVENTION: No. US20020132090A1e1 Nucleic Acid and Protein Homologs  
; FILE REFERENCE: 5800-119  
; CURRENT APPLICATION NUMBER: US/09/833,381  
; CURRENT FILING DATE: 2001-04-11  
; PRIOR APPLICATION NUMBER: 09/516,448  
; PRIOR FILING DATE: 2000-02-29  
; NUMBER OF SEQ ID NOS: 2050  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1147  
; LENGTH: 304  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc.feature  
; LOCATION: (1)..(304)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-833-381-1147

Query Match 6.1%; Score 184.2; DB 9; Length 304;  
Best Local Similarity 79.6%; Pred. No. 9.5e-40;  
Matches 242; Conservative 0; Mismatches 59; Indels 3; Gaps 2;  
Qy 313 TGAAGGGCTCCGCCATCTCGACTGACAAAGATGATTTACAGCC--TTGATTTAGC 370  
Db 304 TAAAGGGCTCCGCCATCTCGACTGACGAGAAATGGTTTCCAGCTTTTGCACCTGCC 245  
Qy 371 AGTTTCAAGGATATGCAAGATTGATCAC-TTCTCTGCTTACAGTGAGCTGATATAC 429  
Db 244 GTTTTCACAGACAGCCGGAAATTTATCATTTTCATCTGTGACACGGGACAGATGTC 185  
Qy 430 AGCAGGTTGATACGGTGGCTCACTGCGCCTCCATATGCTACATAGCTGGCCACCTAG 489

Db 184 AGCAAGTGGGATACGGTGGCCTCACAGCCCTCCACATAGCTCAATAGCTGGACACCCAG 125  
Qy 490 AGGCTGCTGATGCTGCTTGGCAACATGGAGCTAATGTCAATATTCAAGATGCAAGTCTTTT 549  
Db 124 AGGCTGCAAGTGGCTGATGCAACATGGGGCCCAACGTGAATGTTCAAGATGCCGCTTCT 65  
Qy 550 TCACCTCAATGCAATATTGAGCGGTACTATGGACATGAACAGAGTAACTCGCCTTCTTTTGA 609  
Db 64 TCACCCCACTGCACATTTGCGCNTACTATGGGCACGAGCAGGTAAACCAAGTGTCTTTTGA 5  
Qy 610 AATT 613  
Db 4 AGTT 1

## RESULT 13

US-10-424-599-53404  
; Sequence 53404, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Tongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 53404  
; LENGTH: 1953  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_19237C.1  
US-10-424-599-53404

Query Match 3.6%; Score 109.6; DB 13; Length 1953;  
Best Local Similarity 51.3%; Pred. No. 1.3e-18;  
Matches 390; Conservative 0; Mismatches 349; Indels 21; Gaps 5;  
Qy 1452 ATTGGCTCAGGTTCTTTTGGGAAGTATATAAAGGACGATGCAGAAATAAATAGTGGCT 1511  
Db 884 ATTGCATCTGGGTCAATGTGAACCTATTTAAAGGTGATATTTAGCCAGGAAGTGGCC 943  
Qy 1512 ATAAACGTTATCGAGCCCAATACCTACTCTCCAAGTCAGATGTGATATGTTTCCCGA 1571  
Db 944 ATCAAAGTTCTCAAGGCTGACCATGTAAACTCAGAATTCAGAGAGA---GTTCCACAG 1000  
Qy 1572 GAGGTGCCATTTCTGCCAGCTCAATCATCCCTGGTAAATTCAGTTTGTGGTGGCTTGC 1631  
Db 1001 GAAAGTCTATATCATGAGAAGGTTGACACAAGAATGTTGTACAATTCATAGAGCTTGT 1060  
Qy 1632 TTGAATGATCCAGCCAGTTTGGCATTGTCACTCAATACATATCAGGGGTTCTCTGTTTC 1691  
Db 1061 ACCAAGTCCCCACGC---TTATGCTAGTAAAGAAATTTATGCTGTTGGAAGTGTGAT 1117  
Qy 1692 TCCTCTCTCATGAGCAGAGAGATCTTGTATTTGCACTCTAAATTAATTTATTCAGTA 1751  
Db 1118 GACTACCTACATGAAGCAGAGGGCTTTTTTAAATTTCTTACACTGCTCAAAGTAGCAATT 1177  
Qy 1752 GATGTTGCCAAAGGATGGAGTACCTTCACACCTGACACAGCAATATTATCATCGTGAC 1811  
Db 1178 GATGTTTCCAAAGGAATGAATCTTGGACCA-----ACATAATATATCCATAGAGNC 1231  
Qy 1812 TTGAACAGTCACAATATTTCTTCTATAGAGTGGGCATGCTGTGGTGGCAGATTTTGA 1871  
Db 1232 TTGAAGGCTCCCAACCTTTTGTATGGATGAAATTTGACTGTAAAGGTCGCTGATTTGGG 1291  
Qy 1872 GAATCAAGATTTCTACAGTCTCTGGATGAGACCAATGACAAAACACCTCGGACCTC 1931

Db 1292 GTTGTAGAGTTAA-----GCTCAATCTGGGCTCATGAGTCGAGAACTGGAAACATAT 1345  
QY 1932 CATTGATGGCTCTCGAGGTTCACGGAGTGCACTCGGTACACCATCAAGCAGATGTC 1991  
Db 1346 CGATGGATGGCTCTCGAGGTTCATAGAAC-----CAAGCCGTATGATCAACAAGGCTGATGA 1402  
QY 1992 TTCAGTATGCTCTGTCTGTGGAAATTCCTCACTGGCGAAATTCATTTCGCTCATCTC 2051  
Db 1403 TTTAGTTTGGAAATTTTATGGAGTTGCTACCGGAAAGCTTCCATACGATATTTA 1462  
QY 2052 AAGCCAGCGGTGGCGGACGACATGCTTTACCCACATCGAACCTCCATTTGCTAT 2111  
Db 1463 ACCCCCTACAGCGAGCTATAGGAGTGGTTCAAAAGGGTTTGGACCCACCATCCCGAAG 1522  
QY 2112 TCCATTCACAGCCCATATCATCTCTGTATAGAGGTTGGACGAGTCTCTGAGGA 2171  
Db 1523 AACATCTATCCAAAGTTTGTGGAGCTTCTGAGAGGTTTGGACGAGATCCTATATTG 1582  
QY 2172 AGACCCGAATTTCTGAAGTTGTTCATGAAGTTAGAAGT 2211  
Db 1583 AGACCTGATTTCTCGAAATTCATGAGATCCTGCAGCAGT 1622

## RESULT 14

US-10-424-599-129156  
; Sequence 129156, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 129156  
; LENGTH: 3238  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_87633C.1  
US-10-424-599-129156

Query Match 3.6%; Score 107.4; DB 13; Length 3238;  
Best Local Similarity 52.4%; Pred. No. 7.4e-18;  
Matches 387; Conservative 0; Mismatches 331; Indels 21; Gaps 6;  
QY 1453 TTGGCTCAGGTTCTTTGGGAAAGTATATAAGACCATGCGAGAAATAAATAGTGGCTA 1512  
Db 1615 TTGGCTCTGGGTCATTTGGTGACCTGTACAGAGGTACATATTGTAGTCAAGATGTGGCTA 1674  
QY 1513 TAAACGTTATCGAGCCAAATACCTACTGCTCCAGTCAGATGTGGATATGTTTGGCCGAG 1572  
Db 1675 TCAAA-GTCTTAAGCTTGAGGCGATAGTACAGATATGCTGAGG--GAGTTTGCACAGG 1731  
QY 1573 AGGTGTCATCTCTGCGAGCTCAATCATCCCTCGGTAATTCAGTTTGGGGTGTGCTTGGCT 1632  
Db 1732 AAGTTTATATCATGAGGAAGATTCGACACAGAAATGTGTTCAGTTCAITGGGCGCATG-- 1789  
QY 1633 TGAATGATCCAGCCAGTTTCCCATTTGCTCAATACATATCAGGGGGTCTCTGTCTTCT 1692  
Db 1790 -TACTAGGCCCAAACTTTTGTATTTGTTACTGAGTTTATGTTCTAGGGGAGCTTATATG 1848  
QY 1693 CCTCCTTTCATGACGAGAGGATTTCTTGATTTTGCAGTCTAAATTAATTAATTGCGAGTAG 1752  
Db 1849 ACTTTCTGCACAAACAAAGAGGTGTATTTAAGCTTCCATCTTTGCTAAAAGTAGCAATTG 1908  
QY 1753 ATGTTGCCAAAGGATGAGTACCTTCACACCTTGACACGCCAATATATACATCGTGACT 1812  
Db 1909 ATGTTCCAGGGAATGAACCTATTTTGACCA-----AAATATATTAATTCAGGGACC 1962

QY 1813 TGAACAGTCACAATATTTCTTCTATGAGATGGGCATGCTGTGGTGGCAGATTTTGGAG 1872  
Db 1963 TCAAGATGCAATCTTCTGATGGATGAAAATGAAGTGTCAAGTTGCTGATTTTGGG 2022  
QY 1873 AATCAAGATTTTACAGTCTCTCGATGAAGACAAATGACAAACAACTGGGAACCTCC 1932  
Db 2023 TTGCAGGGTGCAA-----ACTCAGTCTGGAGTGATGACAGCTGAACTGGAAACATACC 2076  
QY 1933 GTTGGATGGCTCTCGAGGTTCACGAGTGCACTCGGTACACCATCAAGCAGATGCT 1992  
Db 2077 GTTGGATGGCTCTCGAGG---TCAATGAACACAAACATATGACCAAGGCGAGATGTT 2133  
QY 1993 TCAGTATGCTCTGTCTGTGGGAAATTCCTCACTGGCGAAATTCATTTCGCTCATCTCA 2052  
Db 2134 TCAGTTTGGAAATAGCTCTTTGGAGCTTTTAACTGGAGAACTGCTTACTCTTCCCTGA 2193  
QY 2053 AGCCAGCGGTGGCGGACGACATGGCTTACACACATCAAGCTCCATTTGGCTATT 2112  
Db 2194 CCCAATTACAAGCAGCAGTTGGCGTGGTGCAGAGGGCCTGGGGCTCAATTTCCAAAA 2253  
QY 2113 CCATTCCCAAGCCCATATCATCTCTGTATACGAGGTTGGAACGCATGCTCTGAAGGAA 2172  
Db 2254 ATACGACCCCAAGACTTTCTGAACTTCTTCAGGGTGTGGCAACAAGATCCACGCAA 2313  
QY 2173 GACCCGAATTTCTGAAGT 2191  
Db 2314 GACCAAAATTTCTCTGAAAT 2332

## RESULT 15

US-10-425-114-22955  
; Sequence 22955, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 22955  
; LENGTH: 2207  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: LIB3592-081-E1\_F11  
US-10-425-114-22955

Query Match 3.3%; Score 100.2; DB 13; Length 2207;  
Best Local Similarity 52.1%; Pred. No. 5.4e-16;  
Matches 328; Conservative 0; Mismatches 283; Indels 18; Gaps 4;  
QY 1563 TTTTGGCAGAGGTGTCCATTCTCTCCAGCTCAATCCCTGGGTAATTCAGTTTGTG 1622  
Db 1133 TTTGCTCAGGAAGTATATATTATGAGGAAGTCCCGTCAACAAGATTTGTGTGCAATTTATT 1192  
QY 1623 GGTCTTGTCTGAATGATCCAGCCAGTTTGCAATGTGCATCAATACATATACAGGGGT 1682  
Db 1193 GGGGCTTGCACTAAACCCCAACCTATGT--ATAGTACAGAAATTTATGTCCGCTGA 1249  
QY 1683 TCTCTGTCTCCCTCTTCAATGACGAGAGGATTTCTTGATTTGAGTCTTAATTAATT 1742  
Db 1250 AGTGTGATGATTACCTCCATAAACAAGGTGTTTTTCAAACTTCTCTGTTTAGTGA 1309  
QY 1743 ATTGAGTATGATGTGCGAAGGCATGGAGTACCTTCACAACCTGACACAGCCAATTATA 1802

Db	1310	GTTCACACGGATGTCTCAAAAGCATGAGCTACTTACACAGAAATAT-----ATTATT	1363
Qy	1803	CATCGTGACTTGACAGTACAAATATTCTCTATGAGGATGGGCATGCTGTGGTGGCA	1862
Db	1364	CATCGTGAATTGAAACTGAAACCTTCATGATGAAATGGGACTCTTAAAGTTGCT	1423
Qy	1863	GATTTGGAGAAACAAGATTTCTACAGTCTCTGGATGAAGACAACATGACAAAAACAACCT	1922
Db	1424	GATTTTGGTGTGTGCACGTGTTAAAGCTCAATCTGGAGTT-----ATGACTGCAGAAACT	1477
Qy	1923	GGGAACCTCCGTTGGATGGCTCCTGAGGTGTTCAACGAGTGCATCGGTACACCACTCAA	1982
Db	1478	GGTACTTACCGTTGGATGGCCCGCAGAGGTCATTGACA---CAAACCTATGATCACACAG	1534
Qy	1983	GCAGATCTCTTCAGCTATGCTCTGTGTCTGTGGAAATTTCTACTGGCGAAATTCATTC	2042
Db	1535	GCTGATGTTTTAGTTTTTGGAAATTTTGATGTGGAACTACTCACAGGGAAGATTCCTTAT	1594
Qy	2043	GCTCATCTCAAGCCAGCGGTGGGCAGCAGACATGGCTTACCACCATCAGACTCCC	2102
Db	1595	GAGTACCTGACTCCACTACAAGCGGTGTTGGTGGTGCAGAGGGATTACGGGCTACA	1654
Qy	2103	ATTGGCTATTCCATTCCCAAGCCCATATCATCTCTGTGATACGAGGTGGAAACGCATGT	2162
Db	1655	ATTCCAAAGCATACTCACGCAAGCTTTCGAGCTTCTTCAGAAATGCTGGCAGGAGAC	1714
Qy	2163	CCTGAAGGAAGACCCGAATTTTCTGAAGT	2191
Db	1715	CCTACCCAAAGACCAGACTTCTCCGAAAT	1743

Search completed: September 6, 2004, 11:52:06  
Job time : 922.158 secs

This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 5, 2004, 19:27:21 ; Search time 762.541 Seconds  
(without alignments)  
16852.608 Million cell updates/sec

Title: US-10-626-173-1

Perfect score: 3025

Sequence: 1 gtcaccaccgctccgcc.....aaaaaaaaagggcgccgc 3025

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 337863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_29Jan04.\*

1: Geneseqn1980s.\*

2: Geneseqn1990s.\*

3: Geneseqn2000s.\*

4: Geneseqn2001as.\*

5: Geneseqn2001bs.\*

6: Geneseqn2002s.\*

7: Geneseqn2003as.\*

8: Geneseqn2003bs.\*

9: Geneseqn2003cs.\*

10: Geneseqn2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3025	100.0	3025	3	Aaa47606 Human CAR
2	3025	100.0	3025	7	Acc48580 Human car
3	2508	82.9	2508	4	Aaf44702 Novel pro
4	2505	82.8	2505	3	Aaa47607 Human CAR
5	2458.4	81.3	2788	5	Aas80720 DNA encod
6	2010	66.4	3026	3	Aaa47608 Rat CARDI
7	2010	66.4	3026	7	Acc48581 Rat card
8	1892.2	62.6	2505	3	Aaa47609 Rat CARX
9	1802.4	59.6	2024	8	Aad57334 Human kin
10	466.4	15.4	668	6	Abt09088 Phase-1 R
11	458.4	15.2	633	4	Aas33234 DNA encod
12	432	14.3	640	4	Aas33357 DNA encod
13	135.6	4.5	1557	5	Aas80722 DNA encod
14	106.4	3.5	2283	3	Aac48526 Arabidops
15	100.2	3.3	1704	7	Ada69514 Rice gene
16	98	3.2	1662	6	Abz13209 Arabidops
17	91.4	3.0	3323	9	Adb62159 Human cDN
18	87.8	2.9	1845	7	Ada70011 Rice gene
19	87.4	2.9	1299	4	Aas33288 DNA encod
20	87.4	2.9	1299	5	Abk72043 Human cDN
21	87.4	2.9	1299	5	Abal14324 Human ner
22	87.4	2.9	1299	6	Abk91635 cDNA enco
23	85.6	2.8	3553	9	Adc39121 Novel hum

24	85.6	2.8	4273	4	ABS54855	Abc54855 Human ank
25	85.6	2.8	4519	7	AAD51564	Aad51564 Human str
26	84.6	2.8	1514	3	AAC47638	Aac47638 Arabidops
27	84.2	2.8	5175	6	AAK99407	Aak99407 DNA of AP
28	84	2.8	2915	6	ABN59866	Abn59866 Novel hum
29	84	2.8	13329	5	AA571562	Aas71562 DNA encod
30	84	2.8	14770	6	ABL61893	Abi61893 Colon ade
31	84	2.8	14770	6	ABL68257	Abi68257 Kidney ca
32	84	2.8	14770	6	ABL68538	Abi68538 Kidney ca
33	84	2.8	14770	6	ADB31314	Adb31314 Testoster
34	84	2.8	14797	5	AAS65159	Aas65159 DNA encod
35	84	2.8	15187	5	AAS86131	Aas86131 DNA encod
36	82.6	2.7	6278	4	AAK52293	Aak52293 Human pol
37	80.8	2.7	1797	5	AA50586	Aaf0586 Degenerat
38	80.8	2.7	1797	6	AAD36756	Aad36756 Human RIN
39	80.6	2.7	3160	4	AC90485	Ac90485 Rice EDR1
40	80.2	2.7	883	6	ABK65288	Abk65288 Arabidops
41	80.2	2.7	883	9	ADD30938	Add30938 Plant yie
42	80.2	2.7	883	9	ADS37244	Ade37244 Plant yie
43	79.6	2.6	3438	9	ADB62405	Adb62405 Human cDN
44	79.4	2.6	1314	7	ADA69560	Ada69560 Rice gene
45	79.4	2.6	2187	7	ADA70219	Ada70219 Rice gene

## ALIGNMENTS

RESULT 1  
AAAA47606  
ID AAA47606 standard; cDNA; 3025 BP.

XX AAA47606;

XX 20-OCT-2000 (first entry)

XX Human CARX (Cardiac related Ankyrin-Repeat Protein Kinase) cDNA.

XX Cardiac related ankyrin repeat protein kinase; CARX; cytoskeleton;  
cardiac cell growth factor receptor; cell differentiation; modulator;  
regulator; detection; cellular proliferation; cardiovascular disorder;  
heart failure; hypertension; cancer; sarcoma; ds.

XX Homo sapiens.

XX Key Location/Qualifiers

FT CDS 48..2555

FT /\*tag= a

FT /product= "Human CARX"

XX WO200034330-A1.

XX PD 15-JUN-2000.

XX 10-DEC-1999; 99WO-US029465.

XX 11-DEC-1998; 98US-0111938P.

XX 14-APR-1999; 99US-00291839.

XX (MILL-) MILLENNIUM PHARM INC.

XX Raju J;

XX WPI; 2000-431275/37.

XX P-PSDB; AAB01470.

XX New polynucleotide encoding cardiac-related ankyrin-repeat protein

XX kinase, useful for treating disorders such as cardiovascular disorders,

XX e.g. heart failure and cell differentiation disorders, e.g. cancer.

XX Claim 1; Fig 1; 161pp; English.

XX CARX polypeptides are regulators of signal transduction from cellular

XX receptors (e.g. cardiac cell growth factor receptor) and cytoskeleton

CC function. They also act as modulators of cellular differentiation and  
CC cell death. Nucleic acids, or their fragments encoding CARP polypeptides  
CC are useful for detecting CARP nucleic acids especially mRNA, in a sample.  
CC CARP is useful for treating disorders associated with upregulation or  
CC downregulation of cellular proliferation such as, cardiovascular  
CC disorders (heart failure and hypertension) and disorders associated with  
CC cell differentiation such as cancer and sarcoma

XX Sequence 3025 BP; 882 A; 622 C; 655 G; 866 T; 0 U; 0 Other;

QY 1 GTGACCCACGCTCGGCGCTGAGAGAAAGAAAGAACTTATAATAATGGGAAATTATA 60  
DB 1 GTGACCCACGCTCGGCGCTGAGAGAAAGAAAGAACTTATAATAATGGGAAATTATA 60  
QY 61 AATCTAGACCAACCCAACTTGTACTGATGAATGGAAGAAAAAGTCAGTGAATCATATG 120  
DB 61 AATCTAGACCAACCCAACTTGTACTGATGAATGGAAGAAAAAGTCAGTGAATCATATG 120  
QY 121 TTATCACAATGAAGATTAGAGATGACCTGCGAGTCAAGGAAAAAGAACTGACAGAAC 180  
DB 121 TTATCACAATGAAGATTAGAGATGACCTGCGAGTCAAGGAAAAAGAACTGACAGAAC 180  
QY 181 TAAGGAATATATTTGGCTCTGATGAAGGCTTCAGTAAAGTCAATTTAAATTCGGCACTG 240  
DB 181 TAAGGAATATATTTGGCTCTGATGAAGGCTTCAGTAAAGTCAATTTAAATTCGGCACTG 240  
QY 241 AAAATGGGCTGCTACTTACTTCAATTTATGTTGATTTGGAGGCAAGAAATCATATTC 300  
DB 241 AAAATGGGCTGCTACTTACTTCAATTTATGTTGATTTGGAGGCAAGAAATCATATTC 300  
QY 301 GAACCTTTATGTTGAAGGCTCGGCCATCTGCATGCAAGAAATGGAATTTACAGCCT 360  
DB 301 GAACCTTTATGTTGAAGGCTCGGCCATCTGCATGCAAGAAATGGAATTTACAGCCT 360  
QY 361 TGCAATTTAGCAGTTTACAGGATTAATGCAGAAATGATCACTTCTCTGCTTCAAGTGGAG 420  
DB 361 TGCAATTTAGCAGTTTACAGGATTAATGCAGAAATGATCACTTCTCTGCTTCAAGTGGAG 420  
QY 421 CTGATATACAGAGTTGGATACGGTGGCTCACTGCCCTCCATATTTGTCACAAATAGCTG 480  
DB 421 CTGATATACAGAGTTGGATACGGTGGCTCACTGCCCTCCATATTTGTCACAAATAGCTG 480  
QY 481 GCCACTAGAGCTGCTGATGCTGTTGCAACATGAGGCTTAATGTCAATATTTCAAGATG 540  
DB 481 GCCACTAGAGCTGCTGATGCTGTTGCAACATGAGGCTTAATGTCAATATTTCAAGATG 540  
QY 541 CAGTTTTTTTCACTCCATTTGATGATTAATGCAGGCTATGAGCATGAAGAACTGCGCC 600  
DB 541 CAGTTTTTTTCACTCCATTTGATGATTAATGCAGGCTATGAGCATGAAGAACTGCGCC 600  
QY 601 TTCTTTTCAAAATTTGGTGTGATGAATGTAAGTGTGAAGTGGAGATAGACCCCTCC 660  
DB 601 TTCTTTTCAAAATTTGGTGTGATGAATGTAAGTGTGAAGTGGAGATAGACCCCTCC 660  
QY 661 ACCTAGCATCTGCAAAAGGATTTCTTGAATATTTGCAAACTCTTGTGGAAGAGGCGCA 720  
DB 661 ACCTAGCATCTGCAAAAGGATTTCTTGAATATTTGCAAACTCTTGTGGAAGAGGCGCA 720  
QY 721 AAGCAGATGTAAGTCTCAGATTAATGAAGCACTGTCCTCCACTCCATTTCTGCTCGAT 780  
DB 721 AAGCAGATGTAAGTCTCAGATTAATGAAGCACTGTCCTCCACTCCATTTCTGCTCGAT 780  
QY 781 TTGACACCATGATATAGTTAAGTATCTGCTGCAAAAGTGAATTTGGAAGTTCAACCTCATG 840  
DB 781 TTGACACCATGATATAGTTAAGTATCTGCTGCAAAAGTGAATTTGGAAGTTCAACCTCATG 840  
QY 841 TTGTTAATATCTATGGAGATACCCCTTACACCTGGCATGCTACAAATGGCAATTTGAAG 900  
DB 841 TTGTTAATATCTATGGAGATACCCCTTACACCTGGCATGCTACAAATGGCAATTTGAAG 900

QY 901 TTGCAAGGAATCATCCAAATATCAGAAACAGAAAGTCTGACTAAAGGAAAAACATCTTCA 960  
DB 901 TTGCAAGGAATCATCCAAATATCAGAAACAGAAAGTCTGACTAAAGGAAAAACATCTTCA 960  
QY 961 GTGAAACAGCTTTTCATAGTGTCTGTACCTATGCGAGAGCATTTACCTAGTCAAAATTC 1020  
DB 961 GTGAAACAGCTTTTCATAGTGTCTGTACCTATGCGAGAGCATTTACCTAGTCAAAATTC 1020  
QY 1021 TTCTTGATCAGATGTCTATAAACATCAACCAAGGAGGATGGGACACTGGATTAC 1080  
DB 1021 TTCTTGATCAGATGTCTATAAACATCAACCAAGGAGGATGGGACACTGGATTAC 1080  
QY 1081 ACTCTGCTGTACACAGGTCATATTCGCTGCTGTTTCTGTTTCTTCTGGAATATGAGCTG 1140  
DB 1081 ACTCTGCTGTACACAGGTCATATTCGCTGCTGTTTCTGTTTCTTCTGGAATATGAGCTG 1140  
QY 1141 ATATGAATCTAGTGGCTTGTGATCCGACAGGCTTAGTGGTGAAGGATGAGCAGACAT 1200  
DB 1141 ATATGAATCTAGTGGCTTGTGATCCGACAGGCTTAGTGGTGAAGGATGAGCAGACAT 1200  
QY 1201 GTTTGATGTGGCTTATGAAAAAGGCGATGATGCCATTTGTCACTCTCTGAAAGCATATA 1260  
DB 1201 GTTTGATGTGGCTTATGAAAAAGGCGATGATGCCATTTGTCACTCTCTGAAAGCATATA 1260  
QY 1261 AGACACCACAAGATGAATTCGCCCTGTAATGTAATTTCTCAGCTGAGGAGATGCTCT 1320  
DB 1261 AGACACCACAAGATGAATTCGCCCTGTAATGTAATTTCTCAGCTGAGGAGATGCTCT 1320  
QY 1321 ATGTGCTGTGTTCCATCACTCCCTTGGGGAAGATTAAAAAGCATGACAAAAGAGAGGAGATA 1380  
DB 1321 ATGTGCTGTGTTCCATCACTCCCTTGGGGAAGATTAAAAAGCATGACAAAAGAGAGGAGATA 1380  
QY 1381 TTCTCTCTCTTAAGAGCTGGATTCGCTTCACTTCCATCTTCCAGCTCTCAGAAATGAGT 1440  
DB 1381 TTCTCTCTCTTAAGAGCTGGATTCGCTTCACTTCCATCTTCCAGCTCTCAGAAATGAGT 1440  
QY 1441 TCCATGAGATTTATGGCTCAGGTTCTTTTGGGAAAGTATATAAGGAGCATGACGAAATA 1500  
DB 1441 TCCATGAGATTTATGGCTCAGGTTCTTTTGGGAAAGTATATAAGGAGCATGACGAAATA 1500  
QY 1501 AAATAGTGGCTATATAAGCTTATCGAGCAATACCTACTGCTGCTGCAAGTCAGATGAGATA 1560  
DB 1501 AAATAGTGGCTATATAAGCTTATCGAGCAATACCTACTGCTGCTGCAAGTCAGATGAGATA 1560  
QY 1561 TGTTTTGCCAGAGGTTCCATTTCTGCGAGCTCAATCATCTCCCTGCGTAAATTCAGTTTG 1620  
DB 1561 TGTTTTGCCAGAGGTTCCATTTCTGCGAGCTCAATCATCTCCCTGCGTAAATTCAGTTTG 1620  
QY 1621 TGGTGTCTGCTTGAATGATCCAGGCTTTGCCATTTGCTACTCAATATATATATATATAT 1680  
DB 1621 TGGTGTCTGCTTGAATGATCCAGGCTTTGCCATTTGCTACTCAATATATATATATATAT 1680  
QY 1681 GTTCTGTGTTCTCCCTTCTGATGACAGAGAGATTTCTGATTTGAGTCAATTAATAA 1740  
DB 1681 GTTCTGTGTTCTCCCTTCTGATGACAGAGAGATTTCTGATTTGAGTCAATTAATAA 1740  
QY 1741 TTATTGAGTGTGTTGCCAAAGGATGAGTACTCTTCAACCTGACACAGGCAATTA 1800  
DB 1741 TTATTGAGTGTGTTGCCAAAGGATGAGTACTCTTCAACCTGACACAGGCAATTA 1800  
QY 1801 TACATGCTGATTTGAACAGTCAATATTTCTTCTATGAGATGGGATGCTGCTGGG 1860  
DB 1801 TACATGCTGATTTGAACAGTCAATATTTCTTCTATGAGATGGGATGCTGCTGGG 1860  
QY 1861 CAGATTTTGGAGATCAAGATTTCTACAGTCTCTGATGAGGAGCAACATGACAAACAAAC 1920  
DB 1861 CAGATTTTGGAGATCAAGATTTCTACAGTCTCTGATGAGGAGCAACATGACAAACAAAC 1920  
QY 1921 CTGGGAACCTCGTTGGATGGCTCTGAGGTTTACCGAGTGTGCTCGGTACACCATCA 1980  
DB 1921 CTGGGAACCTCGTTGGATGGCTCTGAGGTTTACCGAGTGTGCTCGGTACACCATCA 1980

QY 1981 AAGCAGATGCTTTCAGCTATGCTCTGTGTGGGAAATTTCTCACTGGCGAAATTCAT 2040  
Db ACC48580  
QY 1981 AAGCAGATGCTTTCAGCTATGCTCTGTGTGGGAAATTTCTCACTGGCGAAATTCAT 2040  
Db ACC48580  
QY 2041 TCGCTCATCTCAAGCCAGCGGCTGGCGAGCAGACATGGCTTACCAACACATCAGACCTC 2100  
Db TCGCTCATCTCAAGCCAGCGGCTGGCGAGCAGACATGGCTTACCAACACATCAGACCTC 2100  
QY 2101 CCATTTGGCTATTCATTCCTCAAGCCCATATCATCTCTGTGTAGAGGTTGGAACGGAT 2160  
Db CCATTTGGCTATTCATTCCTCAAGCCCATATCATCTCTGTGTAGAGGTTGGAACGGAT 2160  
QY 2161 GTCTGAAGGAAGACCCGAAATTTTGAAGTTGTCTATGAAGTTAGAGAGTGTCTCTGCA 2220  
Db GTCTGAAGGAAGACCCGAAATTTTGAAGTTGTCTATGAAGTTAGAGAGTGTCTCTGCA 2220  
QY 2221 ACATTGAGCTGATGCTCTGCTGATCAAGTAACAGAGTGGTCTCTCACCTTCTTCTT 2280  
Db ACATTGAGCTGATGCTCTGCTGATCAAGTAACAGAGTGGTCTCTCACCTTCTTCTT 2280  
QY 2281 CTTCTGATTCGCTGGTGAACCGGGGAGGACCTGGCCGAGTCTATGTGGCAGCATTAAGAA 2340  
Db CTTCTGATTCGCTGGTGAACCGGGGAGGACCTGGCCGAGTCTATGTGGCAGCATTAAGAA 2340  
QY 2341 GTGCTTTCGAATTGGAATATGCTCTAAATGCAAGTCTCTATGCTGTTTGTCCAAAGTG 2400  
Db GTGCTTTCGAATTGGAATATGCTCTAAATGCAAGTCTCTATGCTGTTTGTCCAAAGTG 2400  
QY 2401 CTGGACATATTCCTCTCAAGTCTCTCTTGGAGGAGATGAAGAGTCTTCAATACA 2460  
Db CTGGACATATTCCTCTCAAGTCTCTCTTGGAGGAGATGAAGAGTCTTCAATACA 2460  
QY 2461 CACCAATTGACAAATATGCTATGCTATCCGATCCCATGAGTCAATGCAATTTTCAATCTT 2520  
Db CACCAATTGACAAATATGCTATGCTATCCGATCCCATGAGTCAATGCAATTTTCAATCTT 2520  
QY 2521 GCGGAATAGTACAGCTTTGAGGACAGAGCTGACAGCATTCGGGATATCCTTAAGAG 2580  
Db GCGGAATAGTACAGCTTTGAGGACAGAGCTGACAGCATTCGGGATATCCTTAAGAG 2580  
QY 2581 AGTTTTTTCGCAATGACAGCAAGATTCACCAACGAGCTGGCTTCCAACTATA 2640  
Db AGTTTTTTCGCAATGACAGCAAGATTCACCAACGAGCTGGCTTCCAACTATA 2640  
QY 2641 ACATTTTACTCTAAGGCTCTCTTAAATTTGGGCTGTTTAACTTGCTCTATTAATTC 2700  
Db ACATTTTACTCTAAGGCTCTCTTAAATTTGGGCTGTTTAACTTGCTCTATTAATTC 2700  
QY 2701 CCCACTATTAGCAGCTTTGCAATTTGGCTAAGGAATATATGCAAAAGAACCAAGACA 2760  
Db CCCACTATTAGCAGCTTTGCAATTTGGCTAAGGAATATATGCAAAAGAACCAAGACA 2760  
QY 2761 GAATGTATGAAGAAATGTTTTTAAATTTGTAATTTAAATTTAAATTTAAATTTAAATTT 2820  
Db GAATGTATGAAGAAATGTTTTTAAATTTGTAATTTAAATTTAAATTTAAATTTAAATTT 2820  
QY 2821 TGGAAATGGAGCTTACGCTGCTGGTGGACAGATAAATATGTTTCTCGGCTGAATTT 2880  
Db TGGAAATGGAGCTTACGCTGCTGGTGGACAGATAAATATGTTTCTCGGCTGAATTT 2880  
QY 2881 ATGTAGACTTGTGTTTACAGCTATGGGTTTATTTCTTAGAACATTTGTTCTTTTCTTTT 2940  
Db ATGTAGACTTGTGTTTACAGCTATGGGTTTATTTCTTAGAACATTTGTTCTTTTCTTTT 2940  
QY 2941 CTCATTATGTTTACCTTCTAGTGTTCACCTCTGTGATTAAGATTTCTTTGGTGAATAGAA 3000  
Db CTCATTATGTTTACCTTCTAGTGTTCACCTCTGTGATTAAGATTTCTTTGGTGAATAGAA 3000  
QY 3001 AAAAAAAAAAAAAAGGCGCGCGC 3025  
Db AAAAAAAAAAAAAAGGCGCGCGC 3025

RESULT 2  
ACC48580  
ID ACC48580 standard; cDNA; 3025 BP.  
XX  
AC ACC48580;  
XX  
DT 11-AUG-2003 (first entry)  
XX  
DE Human cardiac-related ankyrin-repeat protein kinase (CARK) cDNA.

XX Human; CARK; cardiac-related ankyrin-repeat protein kinase; enzyme;  
XX cardiant; hypotensive; cytostatic; chromosome 1; gene therapy; gene; ss.  
XX Homo sapiens.

XX Key Location/Qualifiers  
XX CDS 48..2555  
XX /\*tag= a  
XX /product= "Human CARK"  
XX /note= "the CDS is also claimed in Claim 1"  
XX  
XX WO2003020912-A2.  
XX  
XX 13-MAR-2003.  
XX  
XX 04-SEP-2002; 2002WO-US028300.  
XX  
XX 05-SEP-2001; 2001US-00947199.  
XX (MILL-) MILLENNIUM PHARM INC.  
XX  
XX Raju J;  
XX  
XX WPI; 2003-290188/28.

XX Novel isolated cardiac-related ankyrin-repeat protein kinase polypeptide,  
XX useful for treating cellular growth related disorders which include  
XX cardiovascular disorders and proliferative and/or differentiative  
XX disorders.

Claim 1; Fig 1A; 158pp; English.

The present sequence, the coding region of which is also claimed, is that  
of cDNA encoding a novel human cardiac-related ankyrin-repeat protein  
kinase, designated CARK, which plays a role in signalling pathways  
associated with cardiac cell growth or differentiation. The cDNA was  
isolated from a library prepared from subjects suffering from congestive  
heart failure of ischaemic and idiopathic origin. A plasmid containing  
the human CARK nucleotide sequence is deposited as ATCC PTA-1530. The  
CARK gene maps to chromosome 1, within the atrioventricular canal defects  
(AVCD) locus. The gene is therefore a candidate for a congenital heart  
defect susceptibility gene. CARK gene expression is increased in  
ischaemic heart tissue samples, further suggesting a role in the  
regulation of cardiac cell growth and/or differentiation and the  
pathogenesis of cardiovascular disorders, e.g. congestive heart failure  
and cardiac hypertrophy. The invention provides CARK nucleic acids,  
antisense molecules, recombinant expression vectors, host cells and  
transgenic animals in which a CARK gene has been introduced or disrupted,  
and also CARK proteins, fusion proteins, antigenic peptides and anti-CARK  
antibodies. These are useful as modulating agents for regulating a  
variety of cellular processes, e.g. cardiac cellular process, for  
modulating the phosphorylation state of a CARK molecule or one or more  
proteins involved in cellular growth or differentiation, for modulating  
cell behaviour or as targets and therapeutic agents controlling cardiac  
cell proliferation, differentiation, hypertrophy and migration, for  
modulating intra-or inter-cellular signalling and/or gene transcription,  
for modulating cell proliferation, growth, differentiation, survival  
and/or migration, for regulating transmission of signals from cellular  
receptors, for modulating entry of cells, e.g. cardiac precursor cells,  
into mitosis, or for regulating cytoskeletal function. The nucleic acids  
and proteins are useful for treating cellular growth related disorders  
which include cardiovascular disorders (such as heart failure,  
hypertension), and proliferative and/or differentiative disorders (such



QY 1981 AAGCAGATGCTTTCAGCTATGCTCTGTGCTGTGGAAATTCCTCACTGCGAAATTCAT 2040  
Db 1981 AAGCAGATGCTTTCAGCTATGCTCTGTGCTGTGGAAATTCCTCACTGCGAAATTCAT 2040  
QY 2041 TCGCTCATCTCAAGCCAGCGCTGCGCAGCAGCATGCTTACCAACACATCAGACCTC 2100  
Db 2041 TCGCTCATCTCAAGCCAGCGCTGCGCAGCAGCATGCTTACCAACACATCAGACCTC 2100  
QY 2101 CCATTGGCTATTCCATTCCCAAGCCCATATCATCTCTGTGATACGAGGGTGAACGCAT 2160  
Db 2101 CCATTGGCTATTCCATTCCCAAGCCCATATCATCTCTGTGATACGAGGGTGAACGCAT 2160  
QY 2161 GTCTGTGAAGNAGACCCGAATTTCTGAAGTTGTCATGAAGTAGAAGTGTCTCTGCA 2220  
Db 2161 GTCTGTGAAGNAGACCCGAATTTCTGAAGTTGTCATGAAGTAGAAGTGTCTCTGCA 2220  
QY 2221 ACATTGAGCTGATGCTCTCTGCAATCAAGTAACAGCAGTGGTCTCTCTCACCTTCTTCT 2280  
Db 2221 ACATTGAGCTGATGCTCTCTGCAATCAAGTAACAGCAGTGGTCTCTCTCACCTTCTTCT 2280  
QY 2281 CTTCTGATTGCTGTGAACCGGGGAGGACCTGGCGGAGTCAATGGGAGCATTTAAGAA 2340  
Db 2281 CTTCTGATTGCTGTGAACCGGGGAGGACCTGGCGGAGTCAATGGGAGCATTTAAGAA 2340  
QY 2341 GTCTGTTTGAATTTGAATATGCTCTAAATGCAAGGTCCTATGCTGCTTGTCCCAAGTG 2400  
Db 2341 GTCTGTTTGAATTTGAATATGCTCTAAATGCAAGGTCCTATGCTGCTTGTCCCAAGTG 2400  
QY 2401 CTGACAAATATTCCTCAAGGCTGTCTTTGGAGGAGTGAAGAAAGTCTTCAATACA 2460  
Db 2401 CTGACAAATATTCCTCAAGGCTGTCTTTGGAGGAGTGAAGAAAGTCTTCAATACA 2460  
QY 2461 CACCATTGCAATATGCTATGCTATCCGATCCGATGAGCTCAATGCAATTTCACTTCT 2520  
Db 2461 CACCATTGCAATATGCTATGCTATCCGATCCGATGAGCTCAATGCAATTTCACTTCT 2520  
QY 2521 GCCGAATATGAGCTTTGGAGCAGCAGCTGACAGCATTCGGCGTATACCTAAGAG 2580  
Db 2521 GCCGAATATGAGCTTTGGAGCAGCAGCTGACAGCATTCGGCGTATACCTAAGAG 2580  
QY 2581 AGTTTTTCCCGAATGACAGCAACGATTCACACCGCAGCTGGCTTCCAACTATA 2640  
Db 2581 AGTTTTTCCCGAATGACAGCAACGATTCACACCGCAGCTGGCTTCCAACTATA 2640  
QY 2641 ACATTTTACTCTCAAGGCTCTCTTAAATGGGCTGTTTTTACTTGTCTCTATTTAATTC 2700  
Db 2641 ACATTTTACTCTCAAGGCTCTCTTAAATGGGCTGTTTTTACTTGTCTCTATTTAATTC 2700  
QY 2701 CCACTATTAGCAGGCTTCGATTTGTCCTAAGGAATAATATGCAAAAGAACCAAGACA 2760  
Db 2701 CCACTATTAGCAGGCTTCGATTTGTCCTAAGGAATAATATGCAAAAGAACCAAGACA 2760  
QY 2761 GAATGTATGCAAGATTTGTTTTTAAATTTGTAATTAABAAAATTTAGATCGTTACT 2820  
Db 2761 GAATGTATGCAAGATTTGTTTTTAAATTTGTAATTTAAABAAAATTTAGATCGTTACT 2820  
QY 2821 TGGAAATGGAGCCCTAAGCTGTGCTGGACAGATAATAATTATGTTTTCTCGGCTGAAT 2880  
Db 2821 TGGAAATGGAGCCCTAAGCTGTGCTGGACAGATAATAATTATGTTTTCTCGGCTGAAT 2880  
QY 2881 ATGPAGACTGTGTTGACAGCTATGGGTTTATTTCTTAGACATTTGTTCTTTT 2940  
Db 2881 ATGPAGACTGTGTTGACAGCTATGGGTTTATTTCTTAGACATTTGTTCTTTT 2940  
QY 2941 CTCATTATGTTACTTCTPAGTGTTCACCTCTGCTGATTAAGATTTCTTTGTTGAATAGAAA 3000  
Db 2941 CTCATTATGTTACTTCTPAGTGTTCACCTCTGCTGATTAAGATTTCTTTGTTGAATAGAAA 3000  
QY 3001 AAAAAAAAAAAAAAGGGCGGCGC 3025  
Db 3001 AAAAAAAAAAAAAAGGGCGGCGC 3025

RESULT 3  
AAF44702  
ID AAF44702 standard; cDNA; 2508 BP.  
XX  
AC AAF44702;  
XX  
DT 27-MAR-2001 (first entry)  
XX  
DE Novel protein kinase cDNA, SEQ ID NO: 83.  
XX  
KW Human; mouse; protein kinase; antiarthritic; antisclerotic; osteopathic;  
immunopressive; cardiac; renal; antiinflammatory; antiasthmatic;  
dermatological; antidiabetic; antinfertility; gene therapy; vaccine;  
immune disorder; cardiovascular disease; neurodegenerative disease;  
cancer; autoimmune disorder; stroke; inflammatory bowel disease;  
inflammatory pelvic disease; multiple sclerosis; psoriasis; ss.  
XX  
OS Homo sapiens.  
XX  
FN WO200073469-A2.  
XX  
PD 07-DEC-2000.  
XX  
PF 26-MAY-2000; 2000WO-US014842.  
XX  
PR 28-MAY-1999; 99US-0136503P.  
XX  
FA (SUGE-) SUGEN INC.  
XX  
PI Plowman GD, Martinez R, Whyte D, Sudersanam S;  
XX  
DR WPI: 2001-032161/04.  
XX  
DR P-PSDB; AAB65674.  
XX  
PT Nucleic acids encoding kinase polypeptides, useful for diagnosing and  
treating immune-related diseases and disorders, cardiovascular disease,  
neurodegenerative diseases and/or cancers.  
XX  
PS Disclosure; Fig 2; 310pp; English.  
XX  
CC The present sequence encodes a novel protein kinase. The nucleic acids  
and the protein kinases they encode may be used in the treatment and  
diagnosis of diseases associated with inappropriate kinase expression  
such as immune-related diseases and disorders, cardiovascular disease,  
neurodegenerative diseases and/or cancers. The nucleic acids and  
complementary sequences may also be used as DNA probes in diagnostic  
assays. The kinase polypeptides may be used as antigens in the production  
of antibodies of kinase expression and activity. Anti-kinase antibodies  
and kinase antagonists may also be used to down regulate kinase  
expression and activity. Diseases related to kinase expression and  
activity include rheumatoid arthritis, atherosclerosis, autoimmune  
disorders, complications of organ transplantation, myocardial infarction,  
immune disorders, cardiomyopathies, strokes, renal failure, oxidative-  
stress related disorders, chronic inflammatory bowel disease, chronic  
inflammatory pelvic disease, multiple sclerosis, asthma, osteoarthritis,  
psoriasis, rhinitis, autoimmune, diabetes, cancers and reproductive  
disorders  
XX  
SQ Sequence 2508 BP; 722 A; 532 C; 555 G; 699 T; 0 U; 0 Other;  
Query Match 82.9%; Score 2508; DB 4; Length 2508;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2508; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 48 ATGGGAAATATATAATCTAGACCAACCCAACTTGTACTGTGATGAATGGAGAAAAAGTC 107  
Db 1 ATGGGAAATATATAATCTAGACCAACCCAACTTGTACTGTGATGAATGGAGAAAAAGTC 60  
QY 108 AGTGAATCATATCTTATCATCAATAGAAAGATTAGAGATGACCTGCAGATCAAGAAAA 167  
Db 61 AGTGAATCATATCTTATCATCAATAGAAAGATTAGAGATGACCTGCAGATCAAGAAAA 120  
QY 168 GAACCTGACAGAACTAAGGAATATATTGGCTCTGTGAGAGCTTCAGTAAAGTCAATT 227

Db 121 GAACTGACAGAACTAAGGAAATATATTTGGCTCTCATGAAGCCTTCAGTAAAGTCAATTTA 180  
QY 228 AATTACCAGCTGAAATAGGCTGTCTCTACTTCATTTATGTTGCAATTTGTGGAGCAAG 287  
Db 181 AATTACCAGCTGAAATAGGCTGTCTCTACTTCATTTATGTTGCAATTTGTGGAGCAAG 240  
QY 288 AATACACATATTCGAACCTTTATGTTGAAGGGCTCGCCCATCTCGACTGACACAGAAAT 347  
Db 241 AATACACATATTCGAACCTTTATGTTGAAGGGCTCGCCCATCTCGACTGACACAGAAAT 300  
QY 348 GAATTTACAGCCTTCGATTTAGCAGTTTACAGGATAATGCAAGTAATGCAAAATGATCTCTCTG 407  
Db 301 GAATTTACAGCCTTCGATTTAGCAGTTTACAGGATAATGCAAGTAATGATCACTCTCTCTG 360  
QY 408 CTTCAAGTGGAGCTGATATACAGAGGTTGGATACGGTGGCTCACTGCCCTCCATATT 467  
Db 361 CTTCAAGTGGAGCTGATATACAGAGGTTGGATACGGTGGCTCACTGCCCTCCATATT 420  
QY 468 GCTACAAATAGCTGGCCACCTAGAGGCTGCTGATGCTGTTGCAACATGAGCTAATGTC 527  
Db 421 GCTACAAATAGCTGGCCACCTAGAGGCTGCTGATGCTGTTGCAACATGAGCTAATGTC 480  
QY 528 AATATTCAGATGCAATTTTTTCACTCCATTTGATGATGATGAGCGTACTATGACATGAA 587  
Db 481 AATATTCAGATGCAATTTTTTCACTCCATTTGATGATGATGAGCGTACTATGACATGAA 540  
QY 588 CAGGTAACCTCGCTTCTTTGAAATTTGGTGTGATGATGATGATGATGATGATGATGATGAT 647  
Db 541 CAGGTAACCTCGCTTCTTTGAAATTTGGTGTGATGATGATGATGATGATGATGATGATGAT 600  
QY 648 GATAGACCCCTCCACTAGCATCTGCAAAAGGATTTTGAATATTTGCAAACTCTTTGATG 707  
Db 601 GATAGACCCCTCCACTAGCATCTGCAAAAGGATTTTGAATATTTGCAAACTCTTTGATG 660  
QY 708 GAAGAGGAGCAAGAGAGATGTAATGCTCAAGATAATGAAGACCATGTCACCTCCAT 767  
Db 661 GAAGAGGAGCAAGAGAGATGTAATGCTCAAGATAATGAAGACCATGTCACCTCCAT 720  
QY 768 TTCTGTTCTCGATTTGGACACCATGATAGTAAAGTATCTGCTGCAAAAGTATTTGGAA 827  
Db 721 TTCTGTTCTCGATTTGGACACCATGATAGTAAAGTATCTGCTGCAAAAGTATTTGGAA 780  
QY 828 GTTCAACCTCATGTTGTTAAATCTATGGAGATACCCCTTACACCTGGCATGCTACAAT 887  
Db 781 GTTCAACCTCATGTTGTTAAATCTATGGAGATACCCCTTACACCTGGCATGCTACAAT 840  
QY 888 GGCRAATTTGAAGTTGCCAAGGAAATCATCCAAATATCAGGAACAGAAAGTCTGACTAAG 947  
Db 841 GGCRAATTTGAAGTTGCCAAGGAAATCATCCAAATATCAGGAACAGAAAGTCTGACTAAG 900  
QY 948 GAAACATCTTCAGTGAACAGCTTTTATAGTCTTTGATAGTCTGATGCAAGAGCATGAC 1007  
Db 901 GAAACATCTTCAGTGAACAGCTTTTATAGTCTTTGATAGTCTGATGCAAGAGCATGAC 960  
QY 1008 CTAGTCAAAATTTCTTTGATCAGAAATGCTATAAACATCAACCAAGAGGAGGATGG 1067  
Db 961 CTAGTCAAAATTTCTTTGATCAGAAATGCTATAAACATCAACCAAGAGGAGGATGG 1020  
QY 1068 CACACTGGATTACACTCTGCTTGTCTACACGGTCACATTCGCTGTTTCAAGTCTTACTG 1127  
Db 1021 CACACTGGATTACACTCTGCTTGTCTACACGGTCACATTCGCTGTTTCAAGTCTTACTG 1080  
QY 1128 GATAATGAGCTGATATGAATCTAGTGGCTTGTGATCCGAGCAGGCTAGTGTGAAAAA 1187  
Db 1081 GATAATGAGCTGATATGAATCTAGTGGCTTGTGATCCGAGCAGGCTAGTGTGAAAAA 1140  
QY 1188 GATGAGCAGATGTTTCTGATGAGCTTATGAAAGAGGATGATGCAATTTGACACTC 1247  
Db 1141 GATGAGCAGATGTTTCTGATGAGCTTATGAAAGAGGATGATGCAATTTGACACTC 1200  
QY 1248 CTGAAGCATTTAAGAGACCAAGATGAATGGCCTGTAAATGAATTTCTAGCCTGGA 1307

Db 1201 CTGAAGCATTTAAGAGACCAAGATGAATTCGCCCTGTAAATGAATATTTCTCAGCCTGGA 1260  
QY 1308 GGAGATGGCTCTTATGTTGCTGTTCATCACCCCTGGGGAAGATTTAAAGCATGCAAAA 1367  
Db 1261 GGAGATGGCTCTTATGTTGCTGTTCATCACCCCTGGGGAAGATTTAAAGCATGCAAAA 1320  
QY 1368 GAGAAGGCAGATATTCCTCCCTAAAGAGCTGATTCCTTCATTTCCATTTCCATCTTCAGCTC 1427  
Db 1321 GAGAAGGCAGATATTCCTCCCTAAAGAGCTGATTCCTTCATTTCCATTTCCATCTTCAGCTC 1380  
QY 1428 TCAGAAATGAGTTCCATGAGATTTATGGCTCAGGTTCTTTTGGGAAGATATTAAGGA 1487  
Db 1381 TCAGAAATGAGTTCCATGAGATTTATGGCTCAGGTTCTTTTGGGAAGATATTAAGGA 1440  
QY 1488 CGATGCAGAAATAAATAGTGGCTATAAAGAGTATTCGAGCCAAATACCTACTGCTCCCAAG 1547  
Db 1441 CGATGCAGAAATAAATAGTGGCTATAAAGAGTATTCGAGCCAAATACCTACTGCTCCCAAG 1500  
QY 1548 TCAGATGAGATATGTTTTSCGAGAGGTTCCATTTCTCTGCAAGTCAATCATCCCTGC 1607  
Db 1501 TCAGATGAGATATGTTTTSCGAGAGGTTCCATTTCTCTGCAAGTCAATCATCCCTGC 1560  
QY 1608 GTAATTCAGTTTGGGTTGCTTTGAAATGATCCAGCCAGTTTGGCATTGCTCACTCAA 1667  
Db 1561 GTAATTCAGTTTGGGTTGCTTTGAAATGATCCAGCCAGTTTGGCATTGCTCACTCAA 1620  
QY 1668 TACATATCAGGGGTTCTCTGTTCTCCCTCTTCATGAGCAGAGGATTTCTGATTTG 1727  
Db 1621 TACATATCAGGGGTTCTCTGTTCTCCCTCTTCATGAGCAGAGGATTTCTGATTTG 1680  
QY 1728 CAGTCTAAATTAATTTAGTAGATGTTGCCAAGGATGGAGTACCTTCACAACCTG 1787  
Db 1681 CAGTCTAAATTAATTTAGTAGATGTTGCCAAGGATGGAGTACCTTCACAACCTG 1740  
QY 1788 ACAGAGCAATATATACATCGTACTTGAACAGTCAAAATTTCTCTCTATGAGGATGG 1847  
Db 1741 ACAGAGCAATATATACATCGTACTTGAACAGTCAAAATTTCTCTCTATGAGGATGG 1800  
QY 1848 CATCTGTGTGGCAGATTTTGAGAAATCAAGATTTCTACAGTCTCTGGATGAAGACAC 1907  
Db 1801 CATCTGTGTGGCAGATTTTGAGAAATCAAGATTTCTACAGTCTCTGGATGAAGACAC 1860  
QY 1908 ATGACAAACCAACCTGGACCTCCGTTGATGCTGAGTCTTACAGGATTTCTACAGTCTCTG 1967  
Db 1861 ATGACAAACCAACCTGGACCTCCGTTGATGCTGAGTCTTACAGGATTTCTACAGTCTCTG 1920  
QY 1968 CGGTACACCATCAAGAGCAGATGCTTTCAGCTATGCTCTGCTGTGGGAAATTTCTCACT 2027  
Db 1921 CGGTACACCATCAAGAGCAGATGCTTTCAGCTATGCTCTGCTGTGGGAAATTTCTCACT 1980  
QY 2028 GGCAGAAATTCATTCGCTCATCTCAAGCCAGGGCTGGCGAGCAGACATGCTTACCAC 2087  
Db 1981 GGCAGAAATTCATTCGCTCATCTCAAGCCAGGGCTGGCGAGCAGACATGCTTACCAC 2040  
QY 2088 CACATCAGACCTCCCATTTGGCTATTCCATTTCCCAAGCCCATATCATCTCTGCTGATAGA 2147  
Db 2041 CACATCAGACCTCCCATTTGGCTATTCCATTTCCCAAGCCCATATCATCTCTGCTGATAGA 2100  
QY 2148 GGGTGAAGCCGATGCTCTGAAGGAAGACCCGAAATTTCTGGAAGTGTCTGAGATAGAA 2207  
Db 2101 GGGTGAAGCCGATGCTCTGAAGGAAGACCCGAAATTTCTGGAAGTGTCTGAGATAGAA 2160  
QY 2208 GAGTGTCTCTGCAACATTTAGCTGATGTTCTCTGATCAAGTAAAGCAGTGGGCTCTCTC 2267  
Db 2161 GAGTGTCTCTGCAACATTTAGCTGATGTTCTCTGATCAAGTAAAGCAGTGGGCTCTCTC 2220  
QY 2268 TCACCTTCTTCTTCTGATTCCTGTTGAACCCGGGAGGACCTTGGCCGAGTCACTGTC 2327  
Db 2221 TCACCTTCTTCTTCTGATTCCTGTTGAACCCGGGAGGACCTTGGCCGAGTCACTGTC 2280  
QY 2328 GCACATTAAGAGTGTGTTTCAATTTGGAATATGCTCTAAATGCAAGGTCCTATGCTGCT 2387  
Db 2281 GCACATTAAGAGTGTGTTTCAATTTGGAATATGCTCTAAATGCAAGGTCCTATGCTGCT 2340



QY 1128 GATATGGAGCTGATATGAATCTAGTGGCTTGTGATCCCGAGGCTCTAGTGGTGA AAA 1187  
Db 1081 GATAATGGAGCTGATATGAATCTAGTGGCTTGTGATCCCGAGGCTCTAGTGGTGA AAA 1140  
QY 1188 GATGAGCAGACATGTTTGTATGGGCTTATGAAAAGGCGATGATGCCATTTGCACACTC 1247  
Db 1141 GATGAGCAGACATGTTTGTATGGGCTTATGAAAAGGCGATGATGCCATTTGCACACTC 1200  
QY 1248 CTGAAGCATTATAGAGACCAAGATGAATGGCCCTGTATGAATTTCTCAGCCTGGA 1307  
Db 1201 CTGAAGCATTATAGAGACCAAGATGAATGGCCCTGTATGAATTTCTCAGCCTGGA 1260  
QY 1308 GGAGATGGCTCTATGTTGTCTTCCATCACCTTGGGGAAGATTAAAGCATGACAAA 1367  
Db 1261 GGAGATGGCTCTATGTTGTCTTCCATCACCTTGGGGAAGATTAAAGCATGACAAA 1320  
QY 1368 GAGAGGAGATATTTCTCTCTTAAGAGCTGGATGCTTCAATTTCCATCTTCAGCTC 1427  
Db 1321 GAGAGGAGATATTTCTCTCTTAAGAGCTGGATGCTTCAATTTCCATCTTCAGCTC 1380  
QY 1428 TCAGAAATTTGAGTTCATGAGATATTTGGCTCAGTCTTTTGGGAAAGTATATAAGGA 1487  
Db 1381 TCAGAAATTTGAGTTCATGAGATATTTGGCTCAGTCTTTTGGGAAAGTATATAAGGA 1440  
QY 1488 CGATCAGAAATTAATAATAGTGGCTATAAAGCTTATCGAGCCAAATCTACTGCTCAAG 1547  
Db 1441 CGATCAGAAATTAATAATAGTGGCTATAAAGCTTATCGAGCCAAATCTACTGCTCAAG 1500  
QY 1548 TCAGATGTCGATATGTTTGGGAGAGTGTCCATCTCTGCCAGCTCAATCATCCCTGC 1607  
Db 1501 TCAGATGTCGATATGTTTGGGAGAGTGTCCATCTCTGCCAGCTCAATCATCCCTGC 1560  
QY 1608 GTAATTCAGTTTGTGGTGTCTTGAATGATCCAGCCAGTTTGGCATTTGCACTCAA 1667  
Db 1561 GTAATTCAGTTTGTGGTGTCTTGAATGATCCAGCCAGTTTGGCATTTGCACTCAA 1620  
QY 1668 TACATATCAGGGGTTCTGTTCTCCCTCTTATGAGCAGAGAGGATTTGATTG 1727  
Db 1621 TACATATCAGGGGTTCTGTTCTCCCTCTTATGAGCAGAGAGGATTTGATTG 1680  
QY 1728 CAGTCTAAATTAATTTGTCAGTAGATGTTCCAAAGGCATGGAGTACCTTCAACCTG 1787  
Db 1681 CAGTCTAAATTAATTTGTCAGTAGATGTTCCAAAGGCATGGAGTACCTTCAACCTG 1740  
QY 1788 ACACAGCCAAATATACATCGTACCTGAACTGAACAGTCAACAATTTCTTCTATGAGGATGG 1847  
Db 1741 ACACAGCCAAATATACATCGTACCTGAACTGAACAGTCAACAATTTCTTCTATGAGGATGG 1800  
QY 1848 CATGCTGTGGGAGATTTGGAGATCAAGATTTCTACAGTCTCTGGATGAAGCAAC 1907  
Db 1801 CATGCTGTGGGAGATTTGGAGATCAAGATTTCTACAGTCTCTGGATGAAGCAAC 1860  
QY 1908 ATGACAAAACCACTGGAACTCCGTTGGATGGCTCTGAGGTGTTTCAAGCAGTGCAC 1967  
Db 1861 ATGACAAAACCACTGGAACTCCGTTGGATGGCTCTGAGGTGTTTCAAGCAGTGCAC 1920  
QY 1968 CGGTACACCAATCAAGAGATGTTCTCAGCTATGCTGTGTCGTGGGAAATTTCTCACT 2027  
Db 1921 CGGTACACCAATCAAGAGATGTTCTCAGCTATGCTGTGTCGTGGGAAATTTCTCACT 1980  
QY 2028 GCGGAAATTCATTCGCTCATCTCAAGCCAGCGGCTCGGCGAGCAGACATGGTTACCC 2087  
Db 1981 GCGGAAATTCATTCGCTCATCTCAAGCCAGCGGCTCGGCGAGCAGACATGGTTACCC 2040  
QY 2089 CACATCAGACCTCCCATGGCTATTTCCATTCGAGCCATATCATCTCTGCTGATACGA 2147  
Db 2041 CACATCAGACCTCCCATGGCTATTTCCATTCGAGCCATATCATCTCTGCTGATACGA 2100  
QY 2148 GGGTGGAGCGCATGCTCTGAGGAGAGCCCAATTTTCTGAAGTTGTCATGAAGTTAGAA 2207  
Db 2101 GGGTGGAGCGCATGCTCTGAGGAGAGCCCAATTTTCTGAAGTTGTCATGAAGTTAGAA 2160  
QY 2208 GAGTGTCTGTGCAACATTTGAGCTGATGTTCTCTGCTCAAGTAAACGACGCTGGTCTCTC 2267

Db 2161 GAGTGTCTGTGCAACATTTGAGCTGATGTTCTCTCTCATCAAGTAAACAGCAGTGGTCTCTC 2220  
QY 2268 TCACCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTC 2327  
Db 2221 TCACCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTC 2280  
QY 2328 GCAGCATTAAGAGTGGTTCGAAATTTGGAATATGCTCTAAATGCAAGGTCCTATGCTGCT 2387  
Db 2281 GCAGCATTAAGAGTGGTTCGAAATTTGGAATATGCTCTAAATGCAAGGTCCTATGCTGCT 2340  
QY 2388 TTGTCCCAAGTGTGACAAATATTTCTCTCAAGGTCCTCTTTGGAGGAGATGAAAAGA 2447  
Db 2341 TTGTCCCAAGTGTGACAAATATTTCTCTCAAGGTCCTCTTTGGAGGAGATGAAAAGA 2400  
QY 2448 AGTCTTCAATACACACCCATTGACAAATATGCTCTATGCTATGCTATGCTATGCTATGCT 2507  
Db 2401 AGTCTTCAATACACACCCATTGACAAATATGCTCTATGCTATGCTATGCTATGCTATGCT 2460  
QY 2508 CATTTTCATTTCTTCCGAAATAGTAGCAGCTTTTGGAGCAGCAGC 2552  
Db 2461 CATTTTCATTTCTTCCGAAATAGTAGCAGCTTTTGGAGCAGCAGC 2505

## RESULT 5

AAS80720

ID AAS80720 standard; cDNA; 2788 BP.

XX AAS80720;

XX DT 13-FEB-2002 (first entry)

XX DE DNA encoding novel human diagnostic protein #16524.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;

XX KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX OS Homo sapiens.

XX PN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US008631.

XX PR 31-MAR-2000; 2000US-00540217.

XX PR 23-AUG-2000; 2000US-00649167.

XX XX (HYSE-) HYSEQ INC.

XX PI Drmanac RT, Liu C, Tang YT;

XX DR WPI; 2001-639362/73.

XX DR P-PSDB; ABG16533.

XX PT New isolated polynucleotide and encoded polypeptides, useful in

XX PT diagnostics, forensics, gene mapping, identification of mutations

XX PT responsible for genetic disorders or other traits and to assess

XX PT biodiversity.

XX PS Claim 1; SEQ ID NO 16524; 103pp; English.

XX CC The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders

involving aberrant protein expression or biological activity. The  
CC polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic  
CC coding sequences of the invention. Note: The sequence data for this  
CC patent did not appear in the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pat\_sequences  
XX  
SQ Sequence 2788 BP; 807 A; 585 C; 626 G; 770 T; 0 U; 0 Other;

Query Match 81.3%; Score 2458.4; DB 5; Length 2788;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2470; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 85 CTGATGATGAGAGAAAAGTCAGTGAATCATATGTTATCACAATAGAAAAGATTAGAG 144  
DB 317 CTGATGATGAGAGAAAAGTCAGTGAATCATATGTTATCACAATAGAAAAGATTAGAG 376  
QY 145 ATGACCTCGAGATCAAGGAAAAGAACTGACAGAACTTAAGGAATATATTGGCTCTGATG 204  
DB 377 ATGACCTCGGATCAAGGAAAAGAACTGACAGAACTTAAGGAATATATTGGCTCTGATG 436  
QY 205 AAGCCTTCAGTAAAGTCAATTTAAATTAACCGCACTGAAAATGGCTGTCTCTACTTCATT 264  
DB 437 AAGCCTTCAGTAAAGTCAATTTAAATTAACCGCACTGAAAATGGCTGTCTCTACTTCATT 496  
QY 265 TATGTCGATTTGGAGGCAAGAAATCACATTTGCAACTCTTATGTTGAAA- GGGCTC 323  
DB 497 TATGTCGATTTGGAGGCAAGAAATCACATTTGCAACTCTTATGTTGAAAAGGGGCTC 556  
QY 324 CGCCCATCTCGACTGACAAAGAAATGGATTTACAGCCTTGCAATTAGCAGTTTACAAGAT 383  
DB 557 CGCCCATCTCGACTGACAAAGAAATGGATTTACAGCCTTGCAATTAGCAGTTTACAAGAT 616  
QY 384 AATCAGAAATGATCACTTCTGCTTCACAGTGGAGCTGATATACAGAGGTTGGATAC 443  
DB 617 AATCAGAAATGATCACTTCTGCTTCACAGTGGAGCTGATATACAGAGGTTGGATAC 676  
QY 444 GGTGGCTCACTGGCTCCATATTGCTTACAATAGCTGGCCACCTAGAGCTCTGATGTG 503  
DB 677 GGTGGCTCACTGGCTCCATATTGCTTACAATAGCTGGCCACCTAGAGCTCTGATGTG 736  
QY 504 CTGTTGCAACATGAGGCTAATGTCATATTCAGATGTCAGTGTCTTTTCTCACTCCATTCAT 563  
DB 737 CTGTTGCAACATGAGGCTAATGTCATATTCAGATGTCAGTGTCTTTTCTCACTCCATTCAT 796  
QY 564 ATTGCAGCTACTATGACATGACAGTAACTCGCTTCTTTTGAATTTGGTGTGAT 623  
DB 797 ATTGCAGCTACTATGACATGACAGTAACTCGCTTCTTTTGAATTTGGTGTGAT 856  
QY 624 GTAAATGTAAGTGGTGAAGTTGGAGATAGACCCCTCCACCTAGCATCTGCAAAAGGATTC 683  
DB 857 GTAAATGTAAGTGGTGAAGTTGGAGATAGACCCCTCCACCTAGCATCTGCAAAAGGATTC 916  
QY 684 TTGAAATTCGAAAATCTCTTGTATGGAAGAGGAGGAGGAGGAGTGAATGCTCAAGAT 743  
DB 917 TTGAAATTCGAAAATCTCTTGTATGGAAGAGGAGGAGGAGGAGTGAATGCTCAAGAT 976  
QY 744 AATGAAGACCATGTCACCATCCATTTCTGTTCTGATTTGGACACCATCATATAGTTAAG 803  
DB 977 AATGAAGACCATGTCACCATCCATTTCTGTTCTGATTTGGACACCATCATATAGTTAAG 1036  
QY 804 TATCTGTGCAAGTGAATTTGGAGTTCAACCTCATGTTGTTTAATCATATCATGAGATACC 863  
DB 1037 TATCTGTGCAAGTGAATTTGGAGTTCAACCTCATGTTGTTTAATCATATCATGAGATACC 1096  
QY 864 CCCTTACCTGCATGCTACATGCAATTTGAGTTGCCAAGGAATCATCCAAATA 923  
DB 1097 CCCTTACCTGCATGCTACATGCAATTTGAGTTGCCAAGGAATCATCCAAATA 1156

QY 924 TCAGGAACAGAAAGTCTGACTAAGGAAAACATCTTCAGTGAAACAGCTTTTCATAGTGT 983  
DB 1157 TCAGGAACAGAAAGTCTGACTAAGGAAAACATCTTCAGTGAAACAGCTTTTCATAGTGT 1216  
QY 984 TGTACCTATGCGCAAGAGCATTTGACCTAGTCAAAATTTCTTTGATCAGAAATGCTATAAC 1043  
DB 1217 TGTACCTATGCGCAAGAGCATTTGACCTAGTCAAAATTTCTTTGATCAGAAATGCTATAAC 1276  
QY 1044 ATCAACACCAAGAGGAGGATGGGACACCTGGATTAACCTCTGCTGTACACACCGTCAAC 1103  
DB 1277 ATCAACACCAAGAGGAGGATGGGACACCTGGATTAACCTCTGCTGTACACACCGTCAAC 1336  
QY 1104 ATTGCGCTGTTTCAAGTTCTTACTGGATAATGAGAGCTGATATGAATCTAGTGGCTTGTGAT 1163  
DB 1337 ATTGCGCTGTTTCAAGTTCTTACTGGATAATGAGAGCTGATATGAATCTAGTGGCTTGTGAT 1396  
QY 1164 CCAGCAGGCTCTAGTGGTGAAGAAAGATGAGCAGACATGTTTGTATGTTGGCTTATGAAGAA 1223  
DB 1397 CCCAGCAGGCTCTAGTGGTGAAGAAAGATGAGCAGACATGTTTGTATGTTGGCTTATGAAGAA 1456  
QY 1224 GGGCATGATGCCATTTGTACACCTCTGAAAGCATTTATAAGAGACCAAGATGAATGGCC 1283  
DB 1457 GGGCATGATGCCATTTGTACACCTCTGAAAGCATTTATAAGAGACCAAGATGAATGGCC 1516  
QY 1384 TGTAAATGAATTTCTCAGGCTGAGGAGATGGCTCCTATGTTGTCTGTTCCATCACCCCTTG 1343  
DB 1517 TGTAAATGAATTTCTCAGGCTGAGGAGATGGCTCCTATGTTGTCTGTTCCATCACCCCTTG 1576  
QY 1344 GGGAGGATTAAGAGCATGACAAAGAGAGGAGGAGATTTCTCCTCTTAAGAGCTGGATG 1403  
DB 1577 GGGAGGATTAAGAGCATGACAAAGAGAGGAGGAGATTTCTCCTCTTAAGAGCTGGATG 1636  
QY 1404 CTTTCAATTTCCATCTTCCAGCTCTCAGAAATTTGAGTTCCATGAGATTTATGGCTCAGGT 1463  
DB 1637 CTTTCAATTTCCATCTTCCAGCTCTCAGAAATTTGAGTTCCATGAGATTTATGGCTCAGGT 1696  
QY 1464 TCCTTTGGGAAAGTATATAAGACCATGAGAAATAAATAGTGGCTATAAAGCTTAT 1523  
DB 1697 TCCTTTGGGAAAGTATATAAGACCATGAGAAATAAATAGTGGCTATAAAGCTTAT 1756  
QY 1524 CGAGCCAAATACCTACTCTCCAAAGTCAGATGTTGGATATGTTTTCGCGAGAGGTGCCATT 1583  
DB 1757 CGAGCCAAATACCTACTCTCCAAAGTCAGATGTTGGATATGTTTTCGCGAGAGGTGCCATT 1816  
QY 1584 CTCTGCCAGCTCAATCATCTCCGTAATTCAGTTTGTGGGTGCTTGTGTAATGATCC 1643  
DB 1817 CTCTGCCAGCTCAATCATCTCCGTAATTCAGTTTGTGGGTGCTTGTGTAATGATCC 1876  
QY 1644 AGCCAGTTTCCCATTTGTCATCAATACATATCAGGGGGTCTCTGTTCTCCCTCCCTCAT 1703  
DB 1877 AGCCAGTTTCCCATTTGTCATCAATACATATCAGGGGGTCTCTGTTCTCCCTCCCTCAT 1936  
QY 1704 GAGCAGAGAGGATCTTGTATTTGAGTCTAAATTAATTTATTTGCAAGTATGTTGCCAAA 1763  
DB 1937 GAGCAGAGAGGATCTTGTATTTGAGTCTAAATTAATTTATTTGCAAGTATGTTGCCAAA 1996  
QY 1764 GGCATGGATGATCTTCCACACCTGACACAGCCAAATATACATCGTGTGACAGTCAAC 1823  
DB 1997 GGCATGGATGATCTTCCACACCTGACACAGCCAAATATATACATCGTGTGACAGTCAAC 2056  
QY 1824 AATATTTCTTCTATGAGGATGGGATGCTGTGGTGGCAGATTTTGGAGAAATCAAGATTT 1883  
DB 2057 AATATTTCTTCTATGAGGATGGGATGCTGTGGTGGCAGATTTTGGAGAAATCAAGATTT 2116  
QY 1884 CTACAGTCTCTGATGAGAGCAACATGACAAAACAACTGGGAAACCTCGCTTGGATGGCT 1943  
DB 2117 CTACAGTCTCTGATGAGAGCAACATGACAAAACAACTGGGAAACCTCGCTTGGATGGCT 2176  
QY 1944 CTTGAGGTGTTACGAGTGCATCGGTACACCATCAAGCAGATGTTCTTCAAGCTATGCT 2003  
DB 2177 CTTGAGGTGTTACGAGTGCATCGGTACACCATCAAGCAGATGTTCTTCAAGCTATGCT 2236  
QY 2004 CTGTGTTGTGGAAATTTCTCACTGGCGAAATTTCCATTTGGCTCATCTCAAGCCAGCGCT 2063



Db 689 AGGGCTTCTTCAACATTTGTGAAGTCTCTGCTAGAGNAGGAGCAAGCAGATGTGAACG 748  
Qy 736 CTCAGATATGAGACCATGTGCCACTCCATTTCTGTTCTCGATTTGGACACCATGATA 795  
Db 749 CTCAGGACATGAGACACCATGCTCCCTCTGCACTTCTGTTCTCGATTTGGACACCAATA 808  
Qy 796 TAGTAAAGTATCTGCTGCAAGTGAATTTGGAAGTTCAACCTCATGTGTTTATATATCATG 855  
Db 809 TAGTGAGTACCTGCTCCAGAGTGACTTAGAGTCCAGGCTCAGTCAATTAACATCTATG 868  
Qy 856 GAGATACCCCTTACACTGGCATGTACAATGGCAAAATTTGAAGTTGGCAAGGAATCA 915  
Db 869 GTGACACTCTCTTGGCACTGGCATGTACAATGGAAATTTGAAGTTGGCAAGGAATG 928  
Qy 916 TCCAAATATCAGGAACAGAAAGTCTGACTAAGGAAACATCTTCACTGAGAAACAGCTTTTC 975  
Db 929 TCCAGGTAAACAGGAACATGAAAGTCTGACTAAGGAAACATCTTCAAGGAGACAGCTTTTC 988  
Qy 976 ATAGTGTCTTGTACTATGGCAAGACATGCACTAGTCAAAATTTCTTTGATCAGAATG 1035  
Db 989 ACAGTGTCTGTACTATGGCAAGAACATTCACCTGTGCAAAATTTCTTTGATCAGAAATG 1048  
Qy 1036 TCATAACATCAACCAAGGAGGATGGGACACTGGATTAACACTCTGCTTGTCTACC 1095  
Db 1049 CTGTGAACATTAACCAAGGAGGAGATGGGACACAGAGTTGCACTCTGCTTGTCTACC 1108  
Qy 1096 ACAGTCACTTCCGCTGTTTCAAGTCTTCTACTGGAATAGGAGTGAATGAATCTAGTGG 1155  
Db 1109 ACAGCATATCCGCTGTTTCAAGTCTTCTACTTGAATGGTGCAGATATGAATCTTGTG 1168  
Qy 1156 CTGTGTATCCAGAGAGTCTAGTGGTGAAGAAAGATGAGCAGACATGTTTGTGTGGGCTT 1215  
Db 1169 CTGTGTATCCAGCAGTCTAGTGGTGAAGAAAGATGAGCAGACATGTTTGTGTGGGCTT 1228  
Qy 1216 ATGAAAAGGCGATGATGCCATTTGTACACTCTCTGAAGCATATAGAGACACACAAGATG 1275  
Db 1229 ACAGAAAGACATGATGCCATTTGTACACTCTCTGAAGCATATAGAGACACCCAGAGG 1288  
Qy 1276 AATGCCCTGTAAATGAATATTTCTCAGCTGGAGGAGATGGCTCTATGTGTCTGTCCAT 1335  
Db 1289 AGCTGCCATGTAAAGCAATATCCAGCCTGAGGAGATGGCTCTATGTGTCTGTCCCTT 1348  
Qy 1336 CACCTTTGGGAGATTAAGCATGACAAAGAGAGAGGAGAGATTTCTCTTCCCTTAAGAG 1395  
Db 1349 CCCCCCTTGGGCAAGATTAAGCATGACAAAGAGAGGAGATGTTCTCTCTCCCTGAGGG 1408  
Qy 1396 CTGGAATTCCTTCCATTTCCATCTTCAAGTCTCAGAAATTTGAGTTCCTATGAGATATTG 1455  
Db 1409 CTGAACCTACCTCCGCTTCCATCTCCAACTCTCCGAATTCGAGTTCCACGAGATATCG 1468  
Qy 1456 GCTCAGTCTTTTGGAAAGTATATAAGGAGCATGCGAAATATAAATAGTGGCTATAA 1515  
Db 1469 GCTCGGGTCTCTTGGGAAAGTCTATAAGGGGCGATGCGAAATATAAATAGTGGCGATCA 1528  
Qy 1516 AACGTTATCAGGCAATACCTATGCTGCTCAAGTCAGATGTGATATGTTTTCGCGAGAG 1575  
Db 1529 AACGATACCGAGCCACACCTATGCTCCAAAGTCAGACGTGATATGTTTTCGCGAGAG 1588  
Qy 1576 TGTCATTTCTGCGAGCTCAATCACTCCCTGCTGAATTCAGTTTGGGCTGCTCTCTTGA 1635  
Db 1589 TGTCCATTTCTGCGAGCTCAACCAACCCCTCGCTGGTTCAGTTTGGGCTGCTCCCTGG 1648  
Qy 1636 ATGATCCCGGCGATTTGCCATTTGCTACATCAATATACATATCAGGGGTTCTCTGTTCTCC 1695  
Db 1649 ATGACCCGATGATTTGCCATTTGCTACATGATATATATATATATATATATATATATAT 1708  
Qy 1696 TCCTTCATGACGAGAGAGATTTCTGATTTGCAAGTCAAAATTAATTAATTAATTAATTAAT 1755  
Db 1709 TGCTTCATGAACAGAGAGATTTCTGACTTGCAGTCTAAATTAATTAATTAATTAATTAATTA 1768  
Qy 1756 TTGCCAAGGCGATGGAGTACCTTCAACACTGACACAGCCATATATATATATATATATATAT 1815  
Db 1769 TTGCCAAGGCGATGGAGTACCTTCAACAGCTTTGACCCAGGCCAATATATATATATATATAT 1828

Qy 1816 ACAGTCACAATATTTCTTCTATGAGATGGGATGCTGTGTGGCAGATTTTGGAGAT 1875  
Db 1829 ACAGCCACAATATTTCTGCTCTATGAGATGGGCATGCTGTGTGGCAGATTTTGGAGAT 1888  
Qy 1876 CAAGATTTCTACAGTCTCTGGATGAAGACAAATGACAAAAAACAACCTGGGAAACCTCCGTT 1935  
Db 1889 CAAGATTTCTGAGTCCCTGGATGAAGACAAATGACAAAGCAGCCAGGAACTCCGCT 1948  
Qy 1936 GGATGGCTCTGAGGTGTTTCAACGAGTGCACTCGGTACACCATCAAAAGCAGATGTTTCA 1995  
Db 1949 GGATGGCCCTGAGGTGTTTCAACAGTGACGAGATACACCATCAAGGCTGATGTTTCA 2008  
Qy 1996 GCTATGCTCTGCTGTGGGAAATTTCTCACTGGGAAATTTCCATTCGCTCATCTCAAC 2055  
Db 2009 GTTACTCCCTGCTGTGGAGCTCTCACTGGAGAAATTTCCATTCGCTCATCTCAAGC 2068  
Qy 2056 CAGGGCTGGGAGCAGACATGGCTTACCAACATCAGACCTCCCATTTGGCTATTTCA 2115  
Db 2069 CAGCGCTGCAGCAGCAGATATGGCGTATCACACATCAGACCGGCCATCGGCTATTCA 2128  
Qy 2116 TTCCCAAGCCATATCATCTCTGCTGATAGGAGGTGGAAACGATGCTCTGAGGAAGAC 2175  
Db 2129 TCCCAAGCCATCTCATCCCTGCTGATAGGAGGTGGAAATGCAATGCTCTGAGGACGAC 2188  
Qy 2176 CCGAAATTTCTGAAGTTGTCAATGAAGTTAGAAGAGTGTCTCTGCAACATTTGAGCTGATG 2235  
Db 2189 CAGAGTTCTCTGAAGTCGTTAGCAAACTGGAGAGTGCCTATGCAATGTGGAGCTCATGT 2248  
Qy 2236 CTCTGCACTCAAGTAACAGAGTGGTCTCTCTACCTTCTTCTTCTTCTGATTTGCTGG 2295  
Db 2249 CTCAGCATCAAGTAACAGCAGTGGTCTCTCTGCACTTCTCTTCTTCCGATTCGCTGC 2308  
Qy 2296 TGAACCGGGAGACCTGGCCGAGTCATGTGGCAGCATTAAGAAGTCGTTTTCGAATGG 2355  
Db 2309 TGAGCCGGGAGGCGCTGGCCGGAGCCAAGTGGCAGCTTACGAGCGCTTTTGATGG 2368  
Qy 2356 AATATGCTTAAATGCAAGTCTCTATGTGCTTTGTCCTTTCCTTTCCTTCTGCAATATTC 2415  
Db 2369 AGTATGCCCTTAAATGCAAGTCTCTATGCTGGGTGCTCCAAAGTGTGGAAACACACTCTA 2428  
Qy 2416 CTCAAGTCTGCTTTGGAGGAGATGAAGAAGTCTTCAATACACACCATTTGACAAAT 2475  
Db 2429 ATCCGGCCCTGCTTTGGAGGAGATGAATAGGAGACCCAGATTTCACTGTTGACAAAT 2488  
Qy 2476 ATGGCTATGTATCCGATCCCATGAGCTCAATGCAATTTTCAATTTCTTCCGAAATAGTAG 2535  
Db 2489 ACGCTATGTGCTGATCCCATGAGCTGACGACCTTCACTCCCGCAAGACGACAGCA 2548  
Qy 2536 GCTTTGAGGACAGCTGACAGCATTCGGCGTATACCTAAGGAGAGTTCCTCCCGNA 2595  
Db 2549 ACTTTGAGGACAGCACTGACAG-GTCTGGCATACACCTAAGGGGCGTCTCCCATCAGG 2607  
Qy 2596 CTGACAGCAACGATTCACACCGCAAGCTGGCTTCCAACTATATACATTTTACTCTCAA 2655  
Db 2608 CTGACAGCAGTATTTACCCATGGCAGGCTTGTCTTCCAAATTAACGCCCTGCCCTCTG 2667  
Qy 2656 AGGTCTCCTTAAATTTGGGCTTGTGTTTATCTGTGCTATTTTAAATTTCCACATTAAGCAGG 2715  
Db 2668 AGGT-TTCTTCAAAATGCTTGTGTTTATTAAGTCTGTTTAAATTTCCCTTCTACAGGACAG 2726  
Qy 2716 CTTTGGATTTTGGCTTAAGGAATAATATCAAAAGAACCAAGACAGAAATGATATATGAAGA 2775  
Db 2727 GCTTTGACTCATGCC--AAGCCTGAAGTGTCAAGAGCAGATACAGAAATGTCATAGAGA 2784  
Qy 2776 ATTTGTTTTAAATTTGTAATTAATAAAAAAATTTAGATCGTTTACTTGGAAATGGAGCCTA 2835  
Db 2785 ATTTGTTCTTAGTTGATATTTAAAGCCCTTAATGCTGGGCTGGGGTTTCAAA---TC 2840  
Qy 2836 AGTCTGGTGGACAGATAATAATTTATGTTTTTCTGGGCTGAAATATGATAGACTTGTGTT 2895  
Db 2841 TGTGTAGATAGTGGGTTGACCCCTTATGTTTGTAGACCAAACTGTGTGGGCTTGTGTT 2900



Db 749 CTCAGGCAATGAGACCACGTCCTCTCTGACCTTCTGTCTCGATTGGACACCAATA 808  
Qy 796 TAGTTAAGTATCTGCTGCCAAAGTGATTGGAAGTTCAACTCATGTGTTGTAATATCTATG 855  
Db 809 TAGTGAGCTACCTGCTCCAGAGTGACTTAGAGTCCAGGCTCAGCTCATTAACATCTATG 868  
Qy 856 GAGATACCCCTTACACCTCGGCATGTACAAATGGCAAAATTGAAGTTGCCAAGGAAATCA 915  
Db 869 GTGACATCTCTTGGCACTGGCATGCTACATGGAAATTTGNAATTTGCAAGGAAATTG 928  
Qy 916 TCCAAATATCAGGAACAAGAAAGTCTGACTAAGGAAACATCTTTCAGTGAAGAACAGCTTTTC 975  
Db 929 TCCAGGTAACAGGAACTGAAAGTCTGACTAAGGAAACATCTTCAGCGAGACAGCTTTTC 988  
Qy 976 ATAGTGCTTACTTATGGAAGGAGGATTCAGCTTCTGATGATGATGATGATGATGATG 1035  
Db 989 ACAGTGCTTGTACTATGGAAGGAGGATTCAGCTTCTGATGATGATGATGATGATGATG 1048  
Qy 1036 TCATAAACATCAACCAAGGAGGATGGGACACCTGATGATGATGATGATGATGATGATG 1095  
Db 1049 CTGTGAACATTAACCAAGGAGGAGATGGGACACAGGATGACATCTGCTTCTGCTACC 1108  
Qy 1096 ACGGTACATTCGGCTGTTTCAAGTCTTACTGATTAATGGAGTGTATGATGATGATGATG 1155  
Db 1109 ACGGCCATATCCGCTGTTTCAAGTCTTACTTGTATTAATGGTGAATATGATGATGATG 1168  
Qy 1156 CTTGTGATCCAGCAGGTCATGAGTGGTGAAGGAGATGAGCAGACATGTTTGTGATGGGCTT 1215  
Db 1169 CTTGTGATCCAGCAGGTCATGAGTGGTGAAGGAGATGAGCAGACATGTTTGTGATGGGCTT 1228  
Qy 1216 ATGMAAAGGCGATGATGCCATGTCACACTCTCTGAAGCATTAAGAGACCAAGATG 1275  
Db 1229 ACAGAAAGGACATGATGCCATGTTTACACTCTGAAAGCACTACAGAGACCCAGGAGG 1288  
Qy 1276 AATTGCCCTGTAATGATATTTCTCAGCTGGAGGAGATGGCTCCTATGTCGTGTTCCAT 1335  
Db 1289 AGCTGCCATGTACGAATATTTCCAGCTGGAGGAGATGGCTCCTATGTCGTGTTCCCT 1348  
Qy 1336 CAGCCTTGGGAGATTAAGGATGACAAAGAGAGGAGCAGATATTTCTCTCTTAAGAG 1395  
Db 1349 CCGCCTTGGGCAAGATTAAGGATGACAAAGAGAGGAGCAGATTTCTCTCTCTGAGG 1408  
Qy 1396 CTGGATTGCCTTACATTTCCATTTCCAGCTCTCAGAAATGAGTCCATGAGATATTG 1455  
Db 1409 CTGAATACCCCTCCCGCTTCCATCTCCACTCTCCGAATCAGATTCCACGAGATATCG 1468  
Qy 1456 GCTCAGGCTCTTTTGGGAAAGTATATAAGGACGATCGAGAAATAAATAGTGGGTATAA 1515  
Db 1469 GCTCGGCTTCTTTGGGAAAGTCTATAAGGGCGATGCAGAAATAAATAGTGGGATCA 1528  
Qy 1516 AAGTTATCAGGCAATACCTACTGCTCCAGTCCAGTCCAGATGATGATGATGATGATGATG 1575  
Db 1529 AACGATACCGAGCAACACCTACTGCTCCAAGTCCAGCTGATGATGATGATGATGATG 1588  
Qy 1576 TGTCCATTTCTGCGAGTCAATCATCTCCCTCGTAAATTCAGTTTGGGCTGCTGCTTGA 1635  
Db 1589 TGTCCATTTCTGCGAGTCAACCAACCTCGTGTGTTTCAAGTTTGGGCTGCTGCTG 1648  
Qy 1636 ATGATCCAGCAGTGTGCAATGTCATCAATATATATCAGGGGTTCTCTGTTCTCCC 1695  
Db 1649 ATGACCCAGTCAATTTGCAATGTCATCAATATATATCAGGGGTTCTCTGTTCTCCC 1708  
Qy 1696 TCCCTCATGACAGAGAGGATTTGATTTGAGTCAATTAATTAATGAGTATGAGTATGAG 1755  
Db 1709 TGCTTCAAGACAGAGAGGATTTCTTGAATTTGAGTCAATTAATTAATGAGTATGAG 1768  
Qy 1756 TTGCCAAAGGATGGAGTACCTTCAACCTCAGCAGGCAATTAATATGATGATGATGATG 1815  
Db 1769 TTGCCAAAGGATGGAGTACCTCAGCAGCTTGACCCAGCCCAATATACACCCGACCTGA 1828  
Qy 1816 ACAGTCAATATATCTCTATGAGATGGGATGCTGCTGCTGAGATTTTGGAGAT 1875  
Db 1829 ACAGCCAAATATTTCTCTATGAGGATGGCCATGCTGCTGAGGATTTTGGAGAT 1888

Qy 1876 CAAGATTTCTAAGTCTCTGGATGAAGACAAATGACAAACAACTGGGAACCTCGTT 1935  
Db 1889 CAAGATTTCTGAGTCCCTGGATGAAGACAAATGACAAACAAAGCAGGGAACCTCGCT 1948  
Qy 1936 GGATGGCTCTGAGGTTTCAAGCAGTGCCTCGGTACACCATCAAGCAGATGCTTCA 1995  
Db 1949 GGATGGCTCTGAGGTTTCAAGCAGTGCCTCGGTACACCATCAAGCAGATGCTTCA 2008  
Qy 1996 GCTATGCTCTGCTGCTGCTGGGAAATTTCTCACTGGGAAATTTCCATTCGCTCATCTCAAGC 2055  
Db 2009 GTTACTCCCTGCTGCTGCTGGGAGTCTCTCACTGGGAAATTTCCATTCGCTCATCTCAAGC 2068  
Qy 2056 CAGGGCTGGGAGCAGACATGGCTTACCAACATCAGACCTCCCATTTGGCTATTCA 2115  
Db 2069 CAGCGCTGGAGCAGACATATGGCTATCACACATCAGACCGGCCATCGCTATTCA 2128  
Qy 2116 TTCCAAAGCCATATCATCTCTGCTGATACAGGAGTGGAAACCATGCTCTGAAGGAAGAC 2175  
Db 2129 TCCCAAGCCATCTCATCTCTGCTGATACGGGGTGGAAATGCATGCTCTGAAGGACGAC 2188  
Qy 2176 CCGAATTTTCTGAAGTTGTCAAGTTAGAGAGTGTCTCTGCAACATTTGAGCTGATGT 2235  
Db 2189 CAGAGTTCTCTGAAGTCTGTTGCAAACTGGAGAGTGCCTATGCAATGGAGACTCATGT 2248  
Qy 2236 CTCTGCATCAAGTAACAGCAGTGGGTCTCTCTCACTTTCTTTCTTTCTGATTCCTGG 2295  
Db 2249 CTCAGCATCAAGTAACAGCAGTGGGTCTCTCTCACTTTCTTTCTTTCTGATTCCTGGC 2308  
Qy 2296 TGAACGGGAGGACCTGGCCGAGTCATGTGGCAGCATTAAGAAAGTGTGTTTCGAATGG 2355  
Db 2309 TGAGCCGGGAGGCGCTGGCCGAGCAGCTGGCAGCCTTACGGAGCGCTTTTGATTTGG 2368  
Qy 2356 AATATGCTCTAAATGCAAGTCTCTATGCTGTTTGTCCAAAGTGTGCAACATTTCT 2415  
Db 2369 AGTATGCCCTAATGCAAGTCTCTATGCTGGGTGCTCCCAAGTGTGGAAACACTCTA 2428  
Qy 2416 CTCAGGCTCTCTTTGGAGGAGATGAAGAAAGTCTTCAATACACACCATTTGACAAAT 2475  
Db 2429 ATCCGGGCTCTCTTTGGAGGAGATGAATAGGAGCACCAGATTTCAACTGTTGACAAAT 2488  
Qy 2476 ATGGCTATGATCCGATCCATGAGCTCAATGCTTTTCATTTCTGCGAAATAGTACA 2535  
Db 2489 ACGCTATGCTGATCCCATGAGCTGACGACCTTCACTCCCGCCAGACGACAGCA 2548  
Qy 2536 GCTTTGAGGACAGCAGCTGACAGCATTCGGCGTATACCTTAAGGAGAGTTTTTCGCCGAA 2595  
Db 2549 ACITTGAGGACAGCAACTGACAG-GTCTGGCATACACCTAAGGGGCGTCTCCCATCAGG 2607  
Qy 2596 CTGACAGCAAGNTTCCACCAGCGCAAGCTGGCTTCCAACTATTAACATTTTACTCTCA 2655  
Db 2608 CTGACAGCAGTGAATTTACCCATGGCAGCTTGTCTCCAAATTAACGCCCTCGCCCTCTG 2667  
Qy 2656 AGGTCTCTTAAATTTGGGCTTGTTTTACTTGTCTCTTAAATTTAAATTTTAAATTTAGCAGG 2715  
Db 2668 AGGT-TTCITTCAAATGCTCTTGTCTTATTTAGCTCGTTTAAATTTTCTTACAGACAG 2726  
Qy 2716 CTTTGGATTTGGCTTAAGAAATATATGCAAAAGAACCAAGACAGATGATATGAAGA 2775  
Db 2727 GCTTTGACTCATGCC--AAGCCTGAAGTGTCAAGAGCAGATACAGAAATGTCATGAGGA 2784  
Qy 2776 ATTGTTTTTAATTTTCTAAATTAATAAATAAATAATTTAGTCTGTCTTCTGGAATGGAGCCTA 2835  
Db 2785 ATTGTTCTTAGTTTGAATTAAGGCCCTTAATTTGCTGGGCTGGGTTCAAA---TC 2840  
Qy 2836 AGTCTGTGTGGACAGATAAATAATATGTTTTCCTGGGCTGAATTAATGATGATGTTGTT 2895  
Db 2841 TGTGTAGATAGCTGGGTTGACCTTATGTAATTTGTAGACCAAACTGTGTGGGCTTGTGT 2900  
Qy 2896 TGACAG-CTATGGGTTTATTTCTTAGACATGTTCTATTTTCTTCTCATATGTTACT 2954  
Db 2901 TGAGGGTCTCTCTTGGGTTTCTTAAATAAAGAGCTGGCTGATTTTATCTCTCTTCCCTTT 2960

Qy 2955 TCTAGTCTTCACCTCTCTGATTAAAGATCTTTGGTGAATAGAAAAA 3014

nb 2961 ---GTTGTTACTTCTCTGATTAAAGTCTCTTCGGTGATCTAGAAAAA 3016

RESULT 8  
AAA47609  
ID AAA47609 standard; cDNA; 2505 BP.  
XX  
XX  
AAA47609;  
AC  
XX  
20-OCT-2000 (first entry)  
DT  
DT  
XX  
XX  
DE Bat CAPK (Cardiac related Ankyrin-Repeat Protein Kinase) cDNA.

heart failure; hypertension; cancer; sarcoma; ds.  
XX  
OS *Rattus norvegicus*.

Db	61	AGTGAATCTTACGCTATTATCATAGAAAGCGTGGAGGATAACCTGCAGATCAAAAGAAAT	120
Qy	168	GAACTGCAGAGAACTTAAGGAATATATTGGCTCTGATGAAGCCCTTCAGTAAAGTCAATTATA	227
Db	121	GAATTTCAAGAACTTAAGGCAATCTTTGGCTCTGATGAAGCCCTTCAGTGAAGTCAAGTTTA	180
Qy	228	ATTACCGCACTGAAATGGCTGTCTCTACTTTCATTTATGTTGCAATTTGTGGAGGCAAG	287
Db	181	AAATACCGCAGAGAGCTGSCCTGTCCCTGTCTACACCTCTCTGTGTCTGTGGCGCAAC	240
Qy	288	AAATACACATATTGCAACTCTTATGTTGAAAGGGCTCCGCCMATCTCGACTGACAAAGAAAT	347
Db	241	AACTCACATATCCGTGGCCCTTATGTTTAAAGGGCTCCGTCCATCCAGACTGACGAGAAAT	300
Qy	348	GGATTTACAGCTTTGCAATTTAGCAGTTTACAAGGATAATGCAGAAATTCATCACCTTCTCTG	407
Db	301	GGGTTTCAGCTCTGCACTCTGGCCGCTTTACALAGCACGCCCGGAACITTAATCATTCTACTG	360
Qy	408	CTTTCACAGTGAGCTGATATACAGCAGGTTGGATACGGTGGCCTCACTGCCCTCCATATT	467
Db	361	TTGCACAGCGGAGAGATGTTTCAGCAAGTGGGATACGGTGGCCTCACAGCCCTCCACATA	420
Qy	468	GCTCAATAGCTGGCCACCTAGAGGCTGCTGATGCTGTTGCRACATGGAGCTTAAGTGC	527
Db	421	GCTGCAATAGCTGGACACCCAGAGGCTGAGAAGTGTGCTACACAATGGGGCCAAATGTG	480
Qy	528	AAATATCAAGATGCAGATTTTTTTTTCACCTCCATTTGCATATTGCAGCGTACTATGGACATGAA	587
Db	481	AATGTTCAAGATCCGCTCTTTCTTCAACCCACATGCACATTCGACGCTACTATGGCACGAG	540
Qy	588	CAGGTAATCTCGCCTTCTTTTGAATTTTGGTCTGATGTAATGTAAGTGAAGTTGGA	647
Db	541	CAGGTAACAGTGTCTTCTTTGAAGTTTGGTGTGATGTCAATGTAAGCGGTGAAGTTGGG	600
Qy	648	GATGACCCCTCCACCTAGCATCTGCAAAAGGATTTCTTGAATTTGCCAAAACTCTTGATG	707
Db	601	GAAGGCTCTGACACTGGGCTCTGCAAGGGCTTCTTCAACATTTGTGAACTCTCGTA	660
Qy	708	GAAGAAGCGACAAAGCAGATGTGAATGCTCAAGATATGAAGACCATGTGCCACTCCAT	767
Db	661	GAAGAAGCGACAAAGCAGATGTGAACGCTCAGACAATGAAGACCAAGTCCCTCTGCAC	720
Qy	768	TTCTGTTCTCATTTTGACACCATGATAGTTTAAGTATCTGCTGCAAAAGTGAATTGAA	827
Db	721	TTCTGTTCTCATTTTGACACCAATATAGTGAGCTACTGCTCTCAGAGTGAATTAGAG	780
Qy	828	GTTTCAACCTCATGTTGTTTAATATCTATGGAGTACCCCTTACCTCGCATGCTTACAAAT	887
Db	781	GTCCAGCCTCAGCTATTACATCTATGGTGACACTCTTTCACCTCGCATGCTTACAAAT	840
Qy	888	GGCAATTTGAAGTTGCCAGGAATCATCCAATATCAGGAACGAAGTCTGACTAAG	947
Db	841	GGAAATTTTGAAGTTGCCAGGAATTTGTCAGGTAAACAGAACTGAAAGTCTGACTAAG	900
Qy	948	GAAGAACATCTTCAGTGAACAGCTTTTTCATAGTGTCTGTACCTATGGCAAGCAATTGAC	1007
Db	901	GAAGAACATCTTCAGGAGACAGCTTTTTCACAGTGTCTGTACCTATGGCAAGCAATTGAC	960
Qy	1008	CTAGTCAAAATTTCTTTGATCAGATGTCAATAACATCAACCAAGGAGGAGATGGG	1067
Db	961	CTGGTCAAAATTTCTTTGATCAGATGTGTGAAATTAACCAAGGAGGAGATGGG	1020
Qy	1068	CACACTGATTACACTCTGCTTGTCTACACGGTCAATTCGCGCTGGTTCAGTTCCTACTG	1127
Db	1021	CACACAGATTTGCACTCTGCTTGTCTACACGGCCATATCCGCTGGTTCAGTTCCTACTT	1080
Qy	1128	GATAATGAGCTGATATGAATCTAGTGGCTTGTGATCCAGCAGGTCTAGTGGTGAATAA	1187
Db	1081	GATAATGGTCAGATATGAATCTTGTGCTTGTGATCCAGCAGGTCTAGTGGTGAATAA	1140
Qy	1188	GATGAGCAGCATGTTTGAATGTGGGCTTATGAAAAAGGCGATGATGCCATTTGTCACTC	1247

Db 1141 GATGAGCAGACATGTTTGATGTGGGCTTACGAGAAAGGACATGATGCCATTTGTTACACTC 1200  
Qy 1248 CTGAAGCAATTATAGAGACACCAAGATGAATTTGCCCTGTAAATGAATATTTCTCAGCCCTGGA 1307  
Db 1201 CTGAAGCACTACAAGAGACCCAGAGAGAGTGCATGTAACGAATATTTCCAGCCCTGGA 1260  
Qy 1308 GGAGATGCTCTATGTGTCTGTTCCATACACCTTGGGGAAGATTTAAAGCATGACAAAA 1367  
Db 1261 GGAGATGCTCTATGTGTCTGTTCCATACACCTTGGGGAAGATTTAAAGCATGACAAAA 1320  
Qy 1368 GAGAAGGCAGATATTTCTCCTCAAGAGCTGGATTGCTTTCAATTTCCATCTTTGAGCTC 1427  
Db 1321 GAGAAGGCAGATTTCTCCTCAAGAGCTGGATTGCTTTCAATTTCCATCTTTGAGCTC 1380  
Qy 1428 TCAGAAATTGAGTTCCATGAGATTAATTTGGCTCAGTTCTTTTGGGAAAGTATATAAAGGA 1487  
Db 1381 TCCGAAATTCGAGTTCCAGAGATTAATTTGGCTCAGTTCTTTTGGGAAAGTATATAAAGGG 1440  
Qy 1488 CGATTGCAAAATAAATAGTGGCTATAAAGCTTATCGAGCCCAATACCTACTGCTCCAG 1547  
Db 1441 CGATTGCAAAATAAATAGTGGCTATAAAGCTTATCGAGCCCAATACCTACTGCTCCAG 1500  
Qy 1548 TCAGATGTGATATGTTTTGCCAGAGAGTGTCCATTTCTGCGAGCTCAATCATCCCTGC 1607  
Db 1501 TCAGATGTGATATGTTTTGCCAGAGAGTGTCCATTTCTGCGAGCTCAACACCCCTGC 1560  
Qy 1608 GTAAATTCAGTTTGGGTGCTTGTGTTCCATGATCCAGCCAGTTTGCATTTGACTCAA 1667  
Db 1561 GTGTTTCAGTTTGGGTGCTTGTGTTCCATGATCCAGCCAGTTTGCATTTGACTCAA 1620  
Qy 1668 TACATATCAGGGGTTCTCTGTTCTCCTCTCATGAGCAGAGAGGATTTCTGATTTG 1727  
Db 1621 TACATATCAGGGGTTCTCTGTTCTCCTCTCATGAGCAGAGAGATTTCTGATTTG 1680  
Qy 1728 CAGTCTAAATTAATTTGAGTAGATTTGTTGCCAAGGAGTGTGAGTACCTTCAACCTG 1787  
Db 1681 CAGTCTAAATTAATTTGAGTAGATTTGTTGCCAAGGAGTGTGAGTACCTTCAACCTG 1740  
Qy 1788 ACACAGCAATTAATACATGCTGATTTCAACAGTACAAATATTTCTCTATGAGGATGG 1847  
Db 1741 ACCAGCCATATACACCGGACCTGAAAGCAGCAATATTTCTGTTATGAGGATGG 1800  
Qy 1848 CATGCTGTGGTGAGATTTGGAGATTCAGATTTCTCAGTCTCTGGATGAAGCAAC 1907  
Db 1801 CATGCTGTGGTGAGATTTGGAGATTCAGATTTCTCAGTCTCTGGATGAAGCAAC 1860  
Qy 1908 ATGACAAACCACTGGGACCTCCGTTGGATGCTCTGAGGTGTTACGAGTGCAT 1967  
Db 1861 ATGACAAACCACTGGGACCTCCGTTGGATGCTCTGAGGTGTTACGAGTGCAT 1920  
Qy 1968 CGGTACACCATCAAGCAGATGTTCTCAGCTATGCTGTGTGTGGGAAATTTCTCACT 2027  
Db 1921 AGATACCATCAAGCAGATGTTCTCAGCTATGCTGTGTGTGGGAAATTTCTCACT 1980  
Qy 2028 GGGGAAATTCATTCGCTCATCTCAAGCAGCGGCTGGGACAGACATGCGTTACAC 2087  
Db 1981 GGGGAAATTCATTCGCTCATCTCAAGCAGCGGCTGGGACAGACATGCGTTACAC 2040  
Qy 2088 CACATCAGACTCCATTCGCTATTCATTTCCAAAGCCCATATCATCTCTGTTGATAGA 2147  
Db 2041 CACATCAGCGCCCATTCGCTATTCATTTCCAAAGCCCATATCATCTCTGTTGATAGA 2100  
Qy 2148 GGGTGGAAACGATGCTCTGAAGGAAGACCCGAATTTCTGAAGTTGCTATGAAGTTAGAA 2207  
Db 2101 GGGTGGAAATGATGCTCTGAAGGAAGACCCGAATTTCTGAAGTTGCTATGAAGTTAGAA 2160  
Qy 2208 GAGTGTCTCTGCACATTTGAGCTGATGTTCTCTCCATCAAGTAAACAGCAGTGGTCTCTC 2267  
Db 2161 GAGTGTCTCTGCACATTTGAGCTGATGTTCTCTCCATCAAGTAAACAGCAGTGGTCTCTC 2220  
Qy 2268 TCACCTTCTCTCTCTCTGATTTGCTGGTGAACCGGGGAGGACCTGGCGGAGTCTATG 2327  
Db 2221 TCACCTTCTCTCTCTCTGATTTGCTGGTGAACCGGGGAGGACCTGGCGGAGTCTATG 2280

Qy 2328 GCAGCATTAAGAAGTCTGTTTTCGAATTTGAATATGCTCTAAATGCAAGTCTCTATGCTGCT 2387  
Db 2281 GCAGCTTACGGAGCGGTTTTTGATTTGGAGTATGCTTAAATGCAAGTCTCTATGCTGGG 2340  
Qy 2388 TTGTCCCAAAAGTCTGGACAAATATTTCTCTCAAGTCTGTCTTTGGAGGAGATGAAGA 2447  
Db 2341 TGTGTCCTCAAAAGTGTGGACACACTCTAAATCCGGGCTGTCTTTGGAGGAGATGAATAG 2400  
Qy 2448 AGTCTTCAATACACACACCCATTGACAAATATGCTATCCGATCCCATGAGCTCAATG 2507  
Db 2401 AGCACCAGATTTCAACTGTTGACAAATACGGCTATGCTCTGATCCCATGAGCTGACG 2460  
Qy 2508 CATTTTCATTTCTGCCGAAATAGTAGAGCTTTTGAGGACAGCAGC 2552  
Db 2461 CACCTTCACTCCGCAAGCAGACAGCAACTTTTGAGGACAGCAAC 2505

RESULT 9  
AAD57334 standard; cDNA; 2024 BP.  
AC AAD57334;  
XX  
DT 06-NOV-2003 (first entry)  
XX  
DE Human kinase and phosphatase (KPP-7) cDNA.  
XX  
KW Human; kinase; phosphatase; KPP; cell proliferative disorder; hepatitis; atherosclerosis; cirrhosis; haemoglobinuria; polycythaemia vera; cancer; psoriasis; thrombocytopaenia; developmental disorder; Reiter's syndrome; renal tubular acidosis; anaemia; mental retardation; Alzheimer's disease; neurological disorder; Parkinson's disease; epilepsy; glomerulonephritis; autoimmune disorder; inflammatory disorder; AIDS; Goodpasture's syndrome; acquired immune deficiency syndrome; immunosuppressive; Crohn's disease; notropis; transgenic; dermatitis; multiple sclerosis; diabetes mellitus; allergy; gout; Grave's disease; Hashimoto's thyroiditis; bowel syndrome; osteoporosis; rheumatoid arthritis; Sjogren's syndrome; ophthalmological; gene therapy; asthma; anticonvulsant; uropathic; pancreatitis; gene; ss.  
XX Homo sapiens.  
XX  
Key Location/Qualifiers  
FH 46..1857  
FT CDS /\*tag= a  
FT /\*product= "Human KPP"  
XX  
XX WO2003050084-A2.  
XX 19-JUN-2003.  
XX 06-DEC-2002; 2002WO-US039126.  
XX 07-DEC-2001; 2001US-0340235P.  
XX 19-DEC-2001; 2001US-0343007P.  
XX 21-DEC-2001; 2001US-0343546P.  
XX 04-FEB-2002; 2002US-0354388P.  
XX 15-FEB-2002; 2002US-0356757P.  
XX (INCY-) INCYTE GENOMICS INC.  
XX  
XX Kable AE, Chien D, Wilson AD, Swarnakar A, Gorvad AE;  
PI Hafalia AJA, Emerling BM, Ramkumar J, Jin P, Griffin JA, Marquis JP;  
PI Baughn MR, Chawla NK, Lehr-Mason PM, Khare R, Lee S, Hawkins PR;  
PI Becha SD, Lee SY, Sprague WW, Zebbarjadian Y;  
XX  
XX WPI; 2003-532894/50.  
XX P-FSDS; AAE37962.  
XX  
XX New human kinases and phosphatases and polynucleotides, useful for  
PT diagnosing, treating or preventing autoimmune or inflammatory disorders  
PT (e.g. AIDS, allergy or anemia), multiple sclerosis, osteoarthritis,  
PT cancer or hepatitis.

XX PS Claim 5; Page 254; 282bp; English.

XX CC The invention relates to an isolated polypeptide, which is a human kinase

CC and phosphatase (KPP). KPP agonists and antagonists are useful for

CC diagnosing, treating or preventing disorders associated with aberrant

CC expression of KPP, particularly cell proliferative disorders (e.g.

CC arteriosclerosis, atherosclerosis, cirrhosis, hepatitis, paroxysmal

CC nocturnal haemoglobinuria, polycythaemia vera, psoriasis, primary

CC thrombocytopenia or cancer), developmental disorders (eg. renal tubular

CC acidosis, anaemia or mental retardation), neurological disorders (e.g.

CC Alzheimer's disease, Parkinson's disease or epilepsy), autoimmune/

CC inflammatory disorders (e.g. AIDS; acquired immune deficiency syndrome,

CC allergies, asthma, autoimmune thyroiditis, contact dermatitis, Crohn's

CC disease, diabetes mellitus, glomerulonephritis, Goodpasture's syndrome,

CC gout, Grave's disease, Hashimoto's thyroiditis, irritable bowel syndrome,

CC multiple sclerosis, osteoarthritis, osteoporosis, pancreatitis, Reiter's

CC syndrome, rheumatoid arthritis, Sjogren's syndrome, uveitis), or viral,

CC bacterial, fungal, parasitic, protozoan or helminthic infections. The KPP

CC is useful in assessing the effects of exogenous compounds on the

CC expression of nucleic acids and kinases and phosphatases. KPP gene is

CC useful in gene therapy and for creating transgenic animals to model human

CC disease. The present sequence is human KPP cDNA

XX SQ Sequence 2024 BP; 613 A; 431 C; 425 G; 555 T; 0 U; 0 Other;

Query Match 59.6%; Score 1802.4; DB 8; Length 2024;

Best Local Similarity 99.7%; Pred. No. 0;

Matches 1806; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 16 CGGCCCTGAGAAAGGAGAACTTATATAATGGAATTTATAATCTAGACCAACCC 75

DB 14 CTGCCCTGAGAAAGGAGAACTTATATAATGGAATTTATAATCTAGACCAACCC 73

QY 76 AAATCTGTACTGATGAATGGAAGAAAGTCAAGTGAATCATATGTTATCAATAGAAA 135

DB 74 AAATCTGTACTGATGAATGGAAGAAAGTCAAGTGAATCATATGTTATCAATAGAAA 133

QY 136 GATTAGACATGACCTGAGATCAAGGAAAGAAAGTCAAGTGAATCATATGTTATCAATAGAAA 195

DB 134 GATTAGACATGACCTGAGATCAAGGAAAGAAAGTCAAGTGAATCATATGTTATCAATAGAAA 193

QY 196 GCTCTGATGAAGCCCTTCAGTAAAGTCAATTTAAATTCAGGCACTGAAATGGGCTGCTC 255

DB 194 GCTCTGATGAAGCCCTTCAGTAAAGTCAATTTAAATTCAGGCACTGAAATGGGCTGCTC 253

QY 256 TACTTCATTTATGTTGCAATTTGTGAGGCAAGAAATCAATTCGAACTCTTATGTTGA 315

DB 254 TACTTCATTTATGTTGCAATTTGTGAGGCAAGAAATCAATTCGAACTCTTATGTTGA 313

QY 316 AAGGGCTCGGCCCTCTCGACTGACAGAAATGGAATTTACAGCCTTGCAATTTAGCAGTTT 375

DB 314 AAGGGCTCGGCCCTCTCGACTGACAGAAATGGAATTTACAGCCTTGCAATTTAGCAGTTT 373

QY 376 ACAAGATTAATGAGAAATGATCATCTCTGCTTCAGTGGAGCTGATATACAGCAGG 435

DB 374 ACAAGATTAATGAGAAATGATCATCTCTGCTTCAGTGGAGCTGATATACAGCAGG 433

QY 436 TTGGATACGGTGGCCTCACTGCCCTCCATATTTGCTACAAATAGCTGGCCACCTAGAGGCTG 495

DB 434 TTGGATACGGTGGCCTCACTGCCCTCCATATTTGCTACAAATAGCTGGCCACCTAGAGGCTG 493

QY 496 CTGATGCTGTTGCAACATGAGCTAATGTCAATTAATTCAGATGCGATTTTTCATCTC 555

DB 494 CTGATGCTGTTGCAACATGAGCTAATGTCAATTAATTCAGATGCGATTTTTCATCTC 553

QY 556 CATTCGATATTCGAGGCTACTATGACATGACAGTAACCTCGCCTCTCTTTTGAATTTG 615

DB 554 CATTCGATATTCGAGGCTACTATGACATGACAGTAACCTCGCCTCTCTTTTGAATTTG 613

QY 616 GTGCTGATGTAATGTAAGTGGTGAAGTTGGAGATAGACCCCTCCACTAGCATCTGCAA 675

DB 614 GTGCTGATGTAATGTAAGTGGTGAAGTTGGAGATAGACCCCTCCACTAGCATCTGCAA 673

QY 676 AAGGATTTCTGAATATTGCAAAACTCTTGTATGGAAGAGGACGACAAAGCAGATGTGAATG 735

DB 674 AAGGATTTCTGAATATTGCAAAACTCTTGTATGGAAGAGGACGACAAAGCAGATGTGAATG 733

QY 736 CTCGAAGATAATGAAGACCATGTCCCACTCCATTTCTGTTCTCGATTTGGACACCATGATA 795

DB 734 CTCGAAGATAATGAAGACCATGTCCCACTCCATTTCTGTTCTCGATTTGGACACCATGATA 793

QY 796 TAGTTAAGTATCTGCTGCCAAAGTGAATTTGGAAGTTCAACCTCATGTTGTATAATCTATG 855

DB 794 TAGTTAAGTATCTGCTGCCAAAGTGAATTTGGAAGTTCAACCTCATGTTGTATAATCTATG 853

QY 856 GAGATACCCCTTACACCTGGCATGCTACAAATGGCAAAATTTGAAGTTGCCAAGGAATCA 915

DB 854 GAGATACCCCTTACACCTGGCATGCTACAAATGGCAAAATTTGAAGTTGCCAAGGAATCA 913

QY 916 TCCAAATATCAGGAACAGAAAGTCTGACTAAGGAAACATCTTTCAGTGAACAGCTTTTC 975

DB 914 TCCAAATATCAGGAACAGAAAGTCTGACTAAGGAAACATCTTTCAGTGAACAGCTTTTC 973

QY 976 ATAGTGTGTACCTATGGCAAGACATTCACCTAGTCAAAATTTCTTCTGTGATCAGAATG 1035

DB 974 ATAGTGTGTGTACCTATGGCAAGACATTCACCTAGTCAAAATTTCTTCTGTGATCAGAATG 1033

QY 1036 TCATAAAACATCAACCAACCAAGGAGGATGGGCACACTGGATTACACTCTGTTTGTCTACC 1095

DB 1034 TCATAAAACATCAACCAACCAAGGAGGATGGGCACACTGGATTACACTCTGTTTGTCTACC 1093

QY 1096 ACGGTCAATTCGGCTGGTTCAGTCTTACTGTGATTAATGAGCTGATATGAATCTAGTGG 1155

DB 1094 ACGGTCAATTCGGCTGGTTCAGTCTTACTGTGATTAATGAGCTGATATGAATCTAGTGG 1153

QY 1156 CTTGTGATCCAGCAGAGCTAGTGGTGAAGAAAGATGAGCAGACATGTTTGTGATGGGCTT 1215

DB 1154 CTTGTGATCCAGCAGAGCTTGTAGTGGTGAAGAAAGATGAGCAGACATGTTTGTGATGGGCTT 1213

QY 1216 ATGAAAAGGAGCATGATGCCATTTGCACCTCTCTGAAGCATTTAAGACCAACAGATG 1275

DB 1214 ATGAAAAGGAGCATGATGCCATTTGCACCTCTCTGAAGCATTTAAGACCAACAGATG 1273

QY 1276 AATTCGCCCTTAATGAATATTTCTCAGCCTGGAGAGATGGCTCCTATGTTCTGTTCAT 1335

DB 1274 AATTCGCCCTTAATGAATATTTCTCAGCCTGGAGAGATGGCTCCTATGTTCTGTTCAT 1333

QY 1336 CACCTTTGGGGAAGATTTAAAGAGATGACAAAAGAGAGGAGCAGATATTTCTCTCTTAAGAG 1395

DB 1334 CACCTTTGGGGAAGATTTAAAGAGATGACAAAAGAGAGGAGCAGATATTTCTCTCTTAAGAG 1393

QY 1396 CTGATTTGCTTCATTTCCATTTCCAGCTCTCAGAAATTTGATTTCCATGAGATTTATG 1455

DB 1394 CTGATTTGCTTCATTTCCATTTCCAGCTCTCAGAAATTTGATTTCCATGAGATTTATG 1453

QY 1456 GCTCAGGTTCTTTTGGCAAAAGTATATAAAGGACGATGCAAAAATAAATAGTGGCTATAA 1515

DB 1454 GCTCAGGTTCTTTTGGCAAAAGTATATAAAGGACGATGCAAAAATAAATAGTGGCTATAA 1513

QY 1516 AACGTTATTCAGGCAAAATACCTTCTGCTCCAAAGTCAGATGAGGATATTTTGGCCGAGAG 1575

DB 1514 AACGTTATTCAGGCAAAATACCTTCTGCTCCAAAGTCAGATGAGGATATTTTGGCCGAGAG 1573

QY 1576 TGTCCATTTCTGCGACCTCAATCATCTCCCTCGGTAAATTCAGTTTGTGGGTGCTTGTCTGA 1635

DB 1574 TGTCCATTTCTGCGACCTCAATCATCTCCCTCGGTAAATTCAGTTTGTGGGTGCTTGTCTGA 1633

QY 1636 ATGATCCAGCCAGTTTGGCAATTTGCTCACTCAATACATATCAGGGGTTTCTGTCTTCTCC 1695

DB 1634 ATGATCCAGCCAGTTTGGCAATTTGCTCACTCAATACATATCAGGGGTTTCTGTCTTCTCC 1693

QY 1696 TCCCTTCATGACGAAAGAGGATTTCTGATTTGAGTCTAAATTAATTTATTTGCAAGTAGATG 1755

DB 1694 TCCCTTCATGACGAAAGAGGATTTCTGATTTGAGTCTAAATTAATTTATTTGCAAGTAGATG 1753

RESULT 11  
AAS33234  
ID AAS33234 standard; CDNA; 633 BP.  
XX  
AC AAS33234;

XX  
AC

PR 26-JUL-2000; 2000US-0220964P.

Db 197 GGAGATCAAGATTTCGCAGTCCTCGATGAAGACACATGACAAGCAGCCAGGGAAC 256

Db 197 GGAGAAATCAAGATTTCGCAGTCCTCGATGAAGACAAACATGACAAAGCAGCCAGGGAAC 256

PR	14-AUG-2000;	2000US-0224518P.
PR	14-AUG-2000;	2000US-0224519P.
PR	14-AUG-2000;	2000US-0225213P.
PR	14-AUG-2000;	2000US-0225214P.
PR	14-AUG-2000;	2000US-0225266P.
PR	14-AUG-2000;	2000US-0225267P.
PR	14-AUG-2000;	2000US-0225268P.
PR	14-AUG-2000;	2000US-0225270P.
PR	14-AUG-2000;	2000US-0225447P.
PR	14-AUG-2000;	2000US-0225577P.
PR	14-AUG-2000;	2000US-0225758P.
PR	14-AUG-2000;	2000US-0225759P.
PR	18-AUG-2000;	2000US-0226279P.
PR	22-AUG-2000;	2000US-0226681P.
PR	22-AUG-2000;	2000US-0226868P.
PR	23-AUG-2000;	2000US-0227182P.
PR	23-AUG-2000;	2000US-0227009P.
PR	30-AUG-2000;	2000US-0228924P.
PR	01-SEP-2000;	2000US-0229287P.
PR	01-SEP-2000;	2000US-0229343P.
PR	01-SEP-2000;	2000US-0229344P.
PR	01-SEP-2000;	2000US-0229345P.
PR	05-SEP-2000;	2000US-0229509P.
PR	05-SEP-2000;	2000US-0229513P.
PR	06-SEP-2000;	2000US-0230437P.
PR	08-SEP-2000;	2000US-0230438P.
PR	08-SEP-2000;	2000US-0231142P.
PR	08-SEP-2000;	2000US-0231143P.
PR	08-SEP-2000;	2000US-0231144P.
PR	08-SEP-2000;	2000US-0231141P.
PR	08-SEP-2000;	2000US-0232080P.
PR	08-SEP-2000;	2000US-0232081P.
PR	12-SEP-2000;	2000US-0231968P.
PR	14-SEP-2000;	2000US-02323597P.
PR	14-SEP-2000;	2000US-0232398P.
PR	14-SEP-2000;	2000US-0232399P.
PR	14-SEP-2000;	2000US-0232400P.
PR	14-SEP-2000;	2000US-0232401P.
PR	14-SEP-2000;	2000US-0233063P.
PR	14-SEP-2000;	2000US-0233064P.
PR	14-SEP-2000;	2000US-0233065P.
PR	21-SEP-2000;	2000US-0234223P.
PR	21-SEP-2000;	2000US-0234274P.
PR	25-SEP-2000;	2000US-0234997P.
PR	26-SEP-2000;	2000US-0235484P.
PR	27-SEP-2000;	2000US-0235834P.
PR	27-SEP-2000;	2000US-0235836P.
PR	29-SEP-2000;	2000US-0236327P.
PR	29-SEP-2000;	2000US-0236367P.
PR	29-SEP-2000;	2000US-0236368P.
PR	29-SEP-2000;	2000US-0236369P.
PR	29-SEP-2000;	2000US-0236370P.
PR	02-OCT-2000;	2000US-0236802P.
PR	03-OCT-2000;	2000US-0237037P.
PR	03-OCT-2000;	2000US-0237038P.
PR	03-OCT-2000;	2000US-0237039P.
PR	03-OCT-2000;	2000US-0237040P.
PR	13-OCT-2000;	2000US-0239935P.
PR	13-OCT-2000;	2000US-0239937P.
PR	20-OCT-2000;	2000US-0240960P.
PR	20-OCT-2000;	2000US-0241221P.
PR	20-OCT-2000;	2000US-0241785P.
PR	20-OCT-2000;	2000US-0241786P.
PR	20-OCT-2000;	2000US-0241787P.
PR	20-OCT-2000;	2000US-0241808P.
PR	20-OCT-2000;	2000US-0241809P.
PR	20-OCT-2000;	2000US-0241826P.
PR	01-NOV-2000;	2000US-0244617P.
PR	08-NOV-2000;	2000US-0246474P.
PR	08-NOV-2000;	2000US-0246475P.
PR	08-NOV-2000;	2000US-0246476P.
PR	08-NOV-2000;	2000US-0246477P.
PR	08-NOV-2000;	2000US-0246478P.
PR	08-NOV-2000;	2000US-0246523P.
PR	08-NOV-2000;	2000US-0246524P.
PR	08-NOV-2000;	2000US-0246525P.
PR	08-NOV-2000;	2000US-0246526P.
PR	08-NOV-2000;	2000US-0246527P.
PR	08-NOV-2000;	2000US-0246528P.
PR	08-NOV-2000;	2000US-0246532P.
PR	08-NOV-2000;	2000US-0246609P.
PR	08-NOV-2000;	2000US-0246610P.
PR	08-NOV-2000;	2000US-0246611P.
PR	08-NOV-2000;	2000US-0246613P.
PR	17-NOV-2000;	2000US-0249207P.
PR	17-NOV-2000;	2000US-0249208P.
PR	17-NOV-2000;	2000US-0249209P.
PR	17-NOV-2000;	2000US-0249210P.
PR	17-NOV-2000;	2000US-0249211P.
PR	17-NOV-2000;	2000US-0249212P.
PR	17-NOV-2000;	2000US-0249213P.
PR	17-NOV-2000;	2000US-0249214P.
PR	17-NOV-	

CC in samples (e.g. by enzyme linked immunosorbant assay (ELISA)). The  
CC disorders include for example: immune/autoimmune diseases (e.g. HIV  
CC (human immunodeficiency virus) infections, anaemia, rheumatoid arthritis  
CC and multiple sclerosis), cancers and hyperproliferative disorders (e.g.  
CC melanomas, neoplasms of the breast or liver, Sezary syndrome and  
CC Gaucher's disease), neurological diseases (e.g. Alzheimer's disease,  
CC Parkinson's disease and Charcot-Marie-Tooth disease), cardio-  
CC cerebrovascular disorders (e.g. cardiac arrest, tachycardia, angina and  
CC thrombosis), infections caused by bacteria, viruses and fungi and ocular  
CC disorders (e.g. corneal infections). (I) and (II), agonists, antagonists  
CC and antibodies can also be used to promote wound healing, maintain organs  
CC before transplantation, and support cell culture of primary tissues.

Query Match 15.2%; Score 458.4; DB 4; Length 633;  
Best Local Similarity 93.2%; Pred. No. 1.6e-105;  
Matches 517; Conservative 5; Mismatches 27; Indels 6; Gaps 4;  
QY 956 CTTGAGTGAACACGCTTTTCATAGTGGTGTGCTATGCGAAGAGCATTCAGCTAGTCAA 1015  
DB 44 CTTGATGAACACGCTTTTCATAGTGGTGTGCTATGCGAAGAGCATTCAGCTAGTCAA 103  
QY 1016 ATTTCTCTTCATCAGATGTCATAACATCAACACCAAGGAGGATGGGCACACTGG 1075  
DB 104 ATTTCTCTTCATCAGATGTCATAACATCAACACCAAGGAGGATGGGCACACTGG 163  
QY 1076 ATTCACATCTGCTGTGCTACACGGTGCACATCGCCTGGTTCATCTTACTGTGATAATGG 1135  
DB 164 ATTCACATCTGCTGTGCTACACGGTGCACATCGCCTGGTTCATCTTACTGTGATAATGG 223  
QY 1136 AGCTGATATGAATCTAGTGGCTGTGATCCACAGCTCTAGTGGTGAAGAGATGAGCA 1195  
DB 224 AGCTGATATGAATCTAGTGGCTGTGATCCACAGCTCTAGTGGTGAAGAGATGAGCA 283  
QY 1196 GACATGTTGATGTTGGCTTTATGAAAAGGGCATGATGCCATTTGCACACTCTCTGAAGCA 1255  
DB 284 GACATGTTGATGTTGGCTTTATGAAAAGGGCATGATGCCATTTGCACACTCTCTGAAGCA 343  
QY 1256 TTATAAGAGACCAACAGATGATTCGCTGTAATGATATTCACGCTGGAGAGATGG 1315  
DB 344 TTATAAGAGACCAACAGATGATTCGCTGTAATGATATTCACGCTGGAGAGATGG 402  
QY 1316 CTCCTATGTTCTCTTCATCACCCTTGGGGAAGATTAAAGCATCACAAGAGAGAGCG 1375  
DB 403 CTCCTATGTTCTCTTCATCACCCTTGGGGAAGATTAAAGCATCACAAGAGAGAGCG 462  
QY 1376 AGATATTCCTCTCTTAAGAGCTGATGCTTCACATTC - CACTTCAGCTCTCAGAA 1433  
DB 463 AGATATTCCTCTCTTAAGAGCTGATGCTTCACATTCATCTTCAGCTCTTCAGAAA 522  
QY 1434 ATTGAGTTCCATG - AGATATTGGCTCAGTTCTTTTGGGAAGTATATAA - AGGACGA 1490  
DB 523 TTGGATTTCCATGAGGATTATTGGSTCAGTTCTTTTGGGGAAGTATATAAMMGACGR 582  
QY 1491 TGCAGAAATAAATA 1505  
DB 583 TGCWGGAAATTAATA 597

## RESULT 12

AAS33357  
ID AAS33357 standard; cDNA; 640 BP.

XX AC AAS33357;

XX DT 04-DEC-2001 (first entry)

XX DE DNA encoding human secreted protein, Seq ID No 316.

XX KW Immunomodulatory; human immunodeficiency virus; HIV; anaemia; angina;  
XX KW rheumatoid arthritis; antileptosclerotic; cardiac; vascular;  
XX KW cerebroprotective; thrombolytic; antimicrobial; ophthalmological;  
XX KW cytostatic; Alzheimer's disease; Parkinson's disease; human; cancer;  
XX KW multiple sclerosis; cancer; hyperproliferative disorder; infection;

KW Gaucher's disease; neurological disease; cerebrovascular disorder;  
KW thrombosis; wound healing; ss.

XX Homo sapiens.

XX WO200155326-A2.

XX 02-AUG-2001.

XX 17-JAN-2001; 2001WO-US001347.

XX 31-JAN-2000; 2000US-0179065P.

XX 04-FEB-2000; 2000US-0180628P.

XX 24-FEB-2000; 2000US-0184664P.

XX 02-MAR-2000; 2000US-0186350P.

XX 16-MAR-2000; 2000US-0189874P.

XX 17-MAR-2000; 2000US-0190076P.

XX 18-APR-2000; 2000US-0198123P.

XX 19-MAY-2000; 2000US-0205515P.

XX 07-JUN-2000; 2000US-0209467P.

XX 28-JUN-2000; 2000US-0214886P.

XX 30-JUN-2000; 2000US-0215135P.

XX 07-JUL-2000; 2000US-0216547P.

XX 07-JUL-2000; 2000US-0216880P.

XX 11-JUL-2000; 2000US-0217487P.

XX 11-JUL-2000; 2000US-0217496P.

XX 14-JUL-2000; 2000US-0218290P.

XX 26-JUL-2000; 2000US-0220963P.

XX 26-JUL-2000; 2000US-0220964P.

XX 14-AUG-2000; 2000US-0224518P.

XX 14-AUG-2000; 2000US-0224519P.

XX 14-AUG-2000; 2000US-0225213P.

XX 14-AUG-2000; 2000US-0225214P.

XX 14-AUG-2000; 2000US-0225266P.

XX 14-AUG-2000; 2000US-0225267P.

XX 14-AUG-2000; 2000US-0225268P.

XX 14-AUG-2000; 2000US-0225270P.

XX 14-AUG-2000; 2000US-0225477P.

XX 14-AUG-2000; 2000US-0225757P.

XX 14-AUG-2000; 2000US-0225758P.

XX 14-AUG-2000; 2000US-0225759P.

XX 22-AUG-2000; 2000US-0226681P.

XX 22-AUG-2000; 2000US-0226688P.

XX 22-AUG-2000; 2000US-0227182P.

XX 23-AUG-2000; 2000US-0227009P.

XX 30-AUG-2000; 2000US-0228924P.

XX 01-SEP-2000; 2000US-0229287P.

XX 01-SEP-2000; 2000US-0229343P.

XX 01-SEP-2000; 2000US-0229344P.

XX 01-SEP-2000; 2000US-0229345P.

XX 05-SEP-2000; 2000US-0229509P.

XX 06-SEP-2000; 2000US-0229513P.

XX 06-SEP-2000; 2000US-0230437P.

XX 08-SEP-2000; 2000US-0230438P.

XX 08-SEP-2000; 2000US-0231242P.

XX 08-SEP-2000; 2000US-0231243P.

XX 08-SEP-2000; 2000US-0231244P.

XX 08-SEP-2000; 2000US-0231413P.

XX 08-SEP-2000; 2000US-0231414P.

XX 08-SEP-2000; 2000US-0232080P.

XX 12-SEP-2000; 2000US-0232081P.

XX 14-SEP-2000; 2000US-0231968P.

XX 14-SEP-2000; 2000US-0232397P.

XX 14-SEP-2000; 2000US-0232398P.

XX 14-SEP-2000; 2000US-0232399P.

XX 14-SEP-2000; 2000US-0232400P.

XX 14-SEP-2000; 2000US-0232401P.

XX 14-SEP-2000; 2000US-0233063P.

XX 14-SEP-2000; 2000US-0233064P.

XX 14-SEP-2000; 2000US-0233065P.

XX 21-SEP-2000; 2000US-0234223P.

XX 21-SEP-2000; 2000US-0234274P.

```
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239933P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241828P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246533P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 08-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.

PR 05-JAN-2001; 2001US-0259678P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-451931/48.
DR P-PSDB; AAU20648.
XX New nucleic acids and polypeptides, useful for diagnosing, preventing or
PT treating medical conditions.
XX Claim 1; SEQ ID NO 316; 753pp; English.
XX The invention relates to novel isolated nucleic acid molecules (I)
CC encoding human secreted proteins (II). (I) and (II) are used to prevent,
CC treat or ameliorate a medical condition in e.g. humans, mice, rabbits, in
CC goats, horses, cats, dogs, chickens or sheep. (I) and (II) may be used in
CC the prevention, treatment and diagnosis of diseases associated with
CC inappropriate expression of secreted proteins (I) and complementary
CC sequences may also be used as DNA probes in diagnostic assays (e.g.
CC polymerase chain reactions (PCR)) to detect and quantitate the presence
CC of similar nucleic acid sequences in samples, and so which patients may
CC be in need of restorative therapy. (II) may also be used as antigens in
CC the production of antibodies and in assays to identify modulators
CC (agonists and antagonists) of the expression and activity of the secreted
CC proteins. The anti-(II) antibodies and antagonists may also be used to
CC down regulate expression and activity of (II). The anti-(II) antibodies
CC may also be used as diagnostic agents for detecting the presence of (II)
CC in samples (e.g. by enzyme linked immunosorbent assay (ELISA)). The
CC disorders include for example: immune/autoimmune diseases (e.g. HIV
CC (human immunodeficiency virus) infections, anaemia, rheumatoid arthritis
CC and multiple sclerosis), cancers and hyperproliferative disorders (e.g.
CC melanomas, neoplasms of the breast or liver, Sezary syndrome and
CC Gaucher's disease), neurological diseases (e.g. Alzheimer's disease,
CC Parkinson's disease and Charcot-Marie-Tooth disease), cardio-/
CC cerebrovascular disorders (e.g. cardiac arrest, tachycardia, angina and
CC thrombosis), infections caused by bacteria, viruses and fungi and ocular
CC disorders (e.g. corneal infections). (I) and (II), agonists, antagonists
CC and antibodies can also be used to promote wound healing, maintain organs
CC before transplantation, and support cell culture of primary tissues.

Query Match 14.3%; Score 432; DB 4; Length 640;
Best Local Similarity 97.9%; Pred. No. 7.8e-99;
Matches 457; Conservative 1; Mismatches 7; Indels 2; Gaps 2;
QY 956 CTTCACTGAACACAGCTTTTCATAGTGCTTGTACTAGTGGCAAGAGATTGACCTAGTCAA 1015
DB 44 CTTATGAACAGNCCTTTTCATAGTGCTTGTACTAGTGGCAAGAGATTGACCTAGTCAA 103
QY 1016 ATTTCTCTTGATCAGAATGTCATAACATCAACCAAGGAAGGGATGGGCACACTGG 1075
DB 104 ATTTCTCTTGATCAGAATGTCATAACATCAACCAAGGAAGGGATGGGCACACTGG 163
QY 1076 ATTACACTCTGCTTGCTACCAAGGTACATTCGCTGGTTCAGTTCTTACTGGATAATGG 1135
DB 164 ATTACACTCTGCTTGCTACCAAGGTACATTCGCTGGTTCAGTTCTTACTGGATAATGG 223
QY 1136 AGCTGATATGAATCTAGTGGCTTGTGATCCACAGCTCTAGTGGTGAAGAAAGATGAGCA 1195
DB 224 AGCTGATATGAATCTAGTGGCTTGTGATCCACAGCTCTAGTGGTGAAGAAAGATGAGCA 283
QY 1196 GACATGTTGATGGGCTTATGAAAGGGATGATGCCATTGTCACACTCTCTGAAGCA 1255
DB 284 GACATGTTGATGGGCTTATGAAAGGGATGATGCCATTGTCACACTCTCTGAAGCA 343
QY 1256 TTATAAGAGACCAACAGATGAATGCCCTGTATGAATATTCACAGCTGGAGGATGG 1315
DB 344 TTATAAGAGACCAACAGATGAATGCCCTGTATGAATATTCACAGCTGGAGGATGG 402
QY 1316 CTCCTATGTGCTGTTCCATCACCCTTGGGGAAGATTAAGAGATCAACAAAGAGAGGC 1375
DB 403 CTCCTAWGTGCTGTTCCATCACCCTTGGGGAAGATTAAGAGATCAACAAAGAGAGGC 462
```

QY 1376 AGATATTCTCTCTCTAA-GAGCTGGATGCTTTCACATTTCCATCTT 1421  
 |||||  
 Db 463 AGATATTCTCTCTCTAAAGGAGCTGGATGCTTTCACATTTCCATCTT 509  
 |||||

## RESULT 13

AAS80722 ID AAS80722 standard; cDNA; 1557 BP.

AC AAS80722;

DT 13-FEB-2002 (first entry)

DE DNA encoding novel human diagnostic protein #16526.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX Homo sapiens.

PN WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US008631.

XX 31-MAR-2000; 2000US-00540217.

PR 23-AUG-2000; 2000US-00649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

DR P-PSDB; ABG16535.

XX New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity.

XX Claim 1; SEQ ID NO 16526; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)  
 CC sequences. (I) is useful as hybridisation probes, polymerase chain  
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
 CC and in recombinant production of (II). The polynucleotides are also used  
 CC in diagnostics as expressed sequence tags for identifying expressed  
 CC genes. (I) is useful in gene therapy techniques to restore normal  
 CC activity of (II) or to treat disease states involving (II). (II) is  
 CC useful for generating antibodies against it, detecting or quantitating a  
 CC polypeptide in tissue, as molecular weight markers and as a food  
 CC supplement. (II) and its binding partners are useful in medical imaging  
 CC of sites expressing (II). (I) and (II) are useful for treating disorders  
 CC involving aberrant protein expression or biological activity. The  
 CC polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic  
 CC coding sequences of the invention. Note: The sequence data for this  
 CC patent did not appear in the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 1557 BP; 417 A; 400 C; 459 G; 281 T; 0 U; 0 Other;

Query Match 4.5%; Score 135.6; DB 5; Length 1557;

Best Local Similarity 68.4%; Pred. No. 1.1e-23;

Matches 245; Conservative 0; Mismatches 59; Indels 54; Gaps 2;

QY 87 GATGAATGGAAGAAAAAGTCAGTGAATCATATGTTATCACATAGAAAGATTGAAGAT 146

Db 178 GATGAATGGAAGAAAAAGTCAGTGAATCATATGTTATCACATAGAAAGATTGAAGAT 237  
 |||||  
 QY 147 GACCTGCAGATCAAGGAAAAAGAACTGACAGAACTAAAGAAATATATTTGGCTCTGATGAA 206  
 |||||  
 Db 238 GACCTGCAGATCAAGGAAAAAGAACTGACAGAACTAAAGAAATATATTTGGTGGTTTGTG 297  
 |||||  
 QY 207 GCCTT-----CAGTAAA 218  
 |||||  
 Db 298 ATCTTGTGCTTTCAGGAGTGAAGCTGCAGGCGCTTCATGGTGAAGTGTACAGCTCATAAA 357  
 |||||  
 QY 219 GTCAATTTAAATTACCGCACTGAAATGGCTGTCTCTACTTCA-----TTTATGTTGC 272  
 |||||  
 Db 358 GGCAGTCCGACCCCAAGAGTGAGAAAGCAGCAAGATTTATTCGAAAAAGCGAAAGAACAA 417  
 |||||  
 QY 273 ATTTGTGGAGGCAAGAAATCACATATTCGAACCTCTTATGTTGAAAGGGCTCCGCCCATCT 332  
 |||||  
 Db 418 AGCTCCACCGGCAAGAAATCACATATTCGAACCTCTTATGTTGAAAGGGCTCCGCCCATCT 477  
 |||||  
 QY 333 CGACTGACAGAAATGGAATTTACAGCTTGCATTTAGCAGTTTACAGGATAATGCAG 390  
 |||||  
 Db 478 CGACTGACAGAAATGGAATTTACAGCTTGCATTTAGCAGTTTACAGGCGCGAGGAG 535  
 |||||

## RESULT 14

AAC48526

ID AAC48526 standard; DNA; 2283 BP.

AC AAC48526;

DT 18-OCT-2000 (first entry)

XX Arabidopsis thaliana DNA fragment SEQ ID NO: 57797.

XX Hybridisation assay; genetic mapping; gene expression control;  
 KW protein identification; signal transduction pathway; metabolic pathway;  
 KW promoter; termination sequence; ss.

XX Arabidopsis thaliana.

PN EP1033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-00301439.

XX 25-FEB-1999; 99US-0121825P.

PR 05-MAR-1999; 99US-0123180P.

PR 23-MAR-1999; 99US-0123548P.

PR 25-MAR-1999; 99US-0125788P.

PR 29-MAR-1999; 99US-0126264P.

PR 01-APR-1999; 99US-0126785P.

PR 06-APR-1999; 99US-0128234P.

PR 16-APR-1999; 99US-0128714P.

PR 19-APR-1999; 99US-0129845P.

PR 21-APR-1999; 99US-0130077P.

PR 23-APR-1999; 99US-0130510P.

PR 28-APR-1999; 99US-0130891P.

PR 30-APR-1999; 99US-0131449P.

PR 04-MAY-1999; 99US-0132048P.

PR 05-MAY-1999; 99US-0132485P.

PR 06-MAY-1999; 99US-0132486P.

PR 07-MAY-1999; 99US-0132863P.

PR 11-MAY-1999; 99US-0134256P.

PR 14-MAY-1999; 99US-0134218P.

PR 14-MAY-1999; 99US-0134219P.

PR 14-MAY-1999; 99US-0134221P.

PR 14-MAY-1999; 99US-0134370P.



PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161922P.  
PR 28-OCT-1999; 99US-0161933P.  
PR 29-OCT-1999; 99US-0162142P.

Query Match 3.5%; Score 106.4; DB 3; Length 2283;  
Best Local Similarity 52.3%; Pred. No. 3.3e-16;  
Matches 339; Conservative 0; Mismatches 291; Indels 18; Gaps 4;

Qy 1563 TTTTGGCGAGAGTGCCATCTCTCCAGCTCAATCATCCCTCGGTAATTGAGTTTGG 1622  
Db |||||  
Qy 982 TTTTCTCAAGAAGTTTATTAATGAGAAAGTTGACACAAAACGCTGTTCAATTTTG 1041  
Db |||||  
Qy 1623 GGTGCTTGTGATGATGCCAGCAGTTTGGCAATGTCTCAATACATATACAGGGGT 1682  
Db |||||  
Qy 1042 GGTGCATGCACAGATCTCAACCCCTCTGT--ATAGTACTGAGTTTATGCTCGAGGG 1098  
Db |||||  
Qy 1683 TCTCTGTCTCCCTCTCTATGAGCAGAGAGGATCTTGATTTGAGTCTAAATTAAT 1742  
Db |||||  
Qy 1099 AGCATATATGATTTCTTCAACACAGAAATGCGCTTCAAACTTCAAACTTTACTCAA 1158  
Db |||||  
Qy 1743 ATTCCAGTAGATGTGCCAAGGATGAGTACCTTCAACCTGACAGCAATATA 1802  
Db |||||  
Qy 1159 GTTGCACTTGATGCGAAAGAAATGAGTATTTGCATCA-----AAACACATTAAT 1212  
Db |||||  
Qy 1803 CATGCTGACTTGAACTGCAATATTTCTCTATGAGATGGGATGCTGTGGTGA 1862  
Db |||||  
Qy 1213 CACAGGGACCTTAAGACTGCGAATCTTCTTAUGGATGAACATGGACTTGTCAAGTGTCT 1272  
Db |||||  
Qy 1663 GATTTTGGAGAACTCAAGATTTTACAGTCTCTCGATGAAGACAACTCAAAACCACT 1922  
Db |||||  
Qy 1273 GATTTTCGAGTTCAGAGT-----GCAGATTGAATCAGGGTCTATGACTCTGAAACT 1326  
Db |||||  
Qy 1923 GGGAACTCCGTTGGATGCTCTGAGGTGTTTCAACGAGTGCACCTGGTACACCACTCAA 1982  
Db |||||  
Qy 1327 GGGACATACCGTGGATGCTCCAGAGG--TCATTGAGCACAACCTTTACAACTCACAAG 1383  
Db |||||  
Qy 1983 GCAGATGCTTCAAGTATGCTCTGTCTGTGGGAAATTTCTACTGGCAATTCATTC 2042  
Db |||||  
Qy 1384 GCAGATGTGTTCAAGTATGCTATGCTATGGGAACTTCTGACTGGTGACATCCCATAT 1443  
Db |||||  
Qy 2043 GCTCATCTCAAGCCAGCGGCTGGCGAGCAGACATGGCTTACCACCATCAGACTCC 2102  
Db |||||  
Qy 1444 GCTTCTTGACTCCACTACAGCAGCTGTGTGGGTTGTTCCAAAGGGGCTTCGACCCAAA 1503  
Db |||||  
Qy 2103 ATTGGCTATTCCTATCCCAAGCCCATATCATCTCTGCTGATCGAGGGTGGAACTGAT 2162  
Db |||||  
Qy 1504 ATCCCAAAGAAACACACACCCCAAAGTGAAGGGGCTTCTAGAGAGATGCTGGCATCAAGAC 1563  
Db |||||  
Qy 2163 CCTGAAGGAAGACCGGAATTTTCTGAAGTTGTCATGAAGTTAGAGAG 2210  
Db |||||  
Qy 1564 CCAGACAGAGACCACTGTTTGGAGAAATCATAGAAATGCTTCAACAG 1611  
Db |||||

## RESULT 15

ADA69514  
ID ADA69514 standard; DNA; 1704 BP.

AC ADA69514;

XX 20-NOV-2003 (first entry)

DT

XX Rice gene, SEQ ID 2837.

XX Plant; bacterial infection; fungal infection; viral infection; rice;

XX gene; ds.

XX Oryza sativa.

XX OS

XX WO2003000898-A1.

PN

XX 03-JAN-2003.

XX 22-JUN-2001; 2001WO-IB001105.

XX 22-JUN-2001; 2001WO-IB001105.

XX (SYGN) SYNGENTA PARTICIPATIONS AG.

XX Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;

XX Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;

XX WPI; 2003-175290/17.

XX Identifying at least one gene involved in plant resistance or response to

XX pathogenic infection for conferring resistance or tolerance to a plant to

XX bacterial, fungal or viral infection by determining or detecting plant

XX gene expression.

XX Claim 6; SEQ ID NO 2837; 899pp; English.

XX The present invention relates to a method (M1) for identifying genes

XX involved in plant resistance or response to pathogenic infection. M1

XX comprises identifying a gene whose expression is significantly altered in

XX the incompatible interaction of plant gene expression relative to

XX expression of the gene in an uninfected plant, in a mutant plant that

XX does not express a gene associated with response to pathogenic infection,

XX or in a corresponding incompatible or compatible interaction. (M1) is

XX useful for conferring resistance to resistance or tolerance to a plant to

XX bacterial, fungal or viral infection. The present sequence was used to

XX illustrate the invention.

XX Sequence 1704 BP; 463 A; 344 C; 479 G; 417 T; 0 U; 1 Other;

XX Query Match 3.3%; Score 100.2; DB 7; Length 1704;

XX Best Local Similarity 52.1%; Pred. No. 1.1e-14;

XX Matches 328; Conservative 0; Mismatches 283; Indels 18; Gaps 4;

Qy 1563 TTTTGGCGAGAGTGCCATCTCTGCCAGCTCAATCATCCCTGGTAAATTCAGTTTGG 1622

Db |||||

Qy 982 TTTGCCGAGAGTATATATATGAGGAAGTTGCGCAAGAATGTTGTACATTAAT 1041

Db |||||

Qy 1623 GGTGCTTGTGTAATGATCCAGCCAGTTTGGCCATGTGCTCAATACATATCAGGGGT 1682

Db |||||

Qy 1042 GGTGCTATGC---ACTAAACCCCAATCTATGTATGTCACAGATATATGTCAGTGG 1098

Db |||||

Qy 1683 TCTCTGTCTCCCTCTCATGAGCAGAGAGATCTTGATTTGAGTCTAAATTAAT 1742

Db |||||

Qy 1099 AGTGTATTATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAAT 1158

Db |||||

Qy 1743 ATTGCAGTAGATGTTGCCAAAGCATGGAGTACCTTCAACACCTGACACAGCAATATA 1802

Db |||||

Qy 1159 GTTGTGATGATGTTTCAAAAGGCATGAGTACTTGCACCA-----AAACATTAAT 1212

Db |||||

Qy 1803 CATCTGATGCTTGAACAGTCAATATTTCTTCTATGAGGATGGGATGCTGTGGTGA 1862

Db |||||

Qy 1213 CATCGAGATTTGAAAACCTGCTAATCTCTTATGGATGAAAATGGACAGTAAAGTTGCT 1272

Db |||||

Qy 1863 GATTTTGGAGATCAAGATTTCTACAGTCTCTGGATGAAGACAAATGACAAAACAACCT 1922

Db |||||

Qy 1273 GATTTTGGTGTGCACGTTTAAAGCTCAATCTGAGTA-----ATGACTGTGAACA 1326

Db |||||

Qy 1923 GGGAACTCCGTTGGATGCTCTGAGGTGTTTCAACGAGTGCACCTCGGTACACCACTAAA 1982

Db |||||

Qy 1327 GGGACATATCGTTGGATGGCTCCAGAGGTTATAGAACA---CAAGCCCTATGATCACAAG 1383

Db |||||

Qy 1983 GCAGATGCTTCAAGTATGCTCTGTGCTGTGGGAAATTTCTACTGGGAAATTCCTATTC 2042

Db |||||

Qy 1384 GCGGATGTTTAAAGTTTGGAAATTTGATGTTGGGAGTCTGCTACAGGAAATTCCTTAT 1443

Db |||||

Qy 2043 GCTCATCTCAAGCCAGCGGCTGCGGAGCAGACATGGCTTACCACCATCAGACTCCC 2102

Db |||||

Qy 1444 GAGTACCTGACACCATTAACAAGCAGCTGTAGTGTGGTTCAGAGGGATTACGGCCTACA 1503

Db |||||

Tue Sep 7 12:10:51 2004

Qy	2103	ATTGGCTATTCCATTCCCAAGCCCATATCATCTCTGCTGATACGAGGTGGAAACGCATGT	2162
Db	1504	ATACCAAAAATGCTCATGCAAAACTTTCTGAGCTTCTTCAGAAATGTTGGCAACAGGAG	1563
Qy	2163	CCTGAAGGAAGACCCGAATTTTCTGAAGT	2191
Db	1564	CCTGCTGAAAGACCAGATTTCTCTGAAAT	1592

Search completed: September 6, 2004, 01:05:49  
Job time : 771.541 secs

score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

OM nucleic - nucleic search, using sw model  
Run on: September 5, 2004, 19:33:36 ; Search time 7778.57 Seconds  
(without alignments)  
16855.624 Million cell updates/sec

Title: US-10-626-173-1  
Perfect score: 3025  
Sequence: 1 gtcgaccacgctcgccgcccc.....aaaaaaaaagggcgccgc 3025

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl.\*

- 1: gb.ba.\*
- 2: gb.htg.\*
- 3: gb.in.\*
- 4: gb.om.\*
- 5: gb.ov.\*
- 6: gb.pat.\*
- 7: gb.ph.\*
- 8: gb.pl.\*
- 9: gb.pr.\*
- 10: gb.ro.\*
- 11: gb.sts.\*
- 12: gb.sy.\*
- 13: gb.un.\*
- 14: gb.vi.\*
- 15: em.ba.\*
- 16: em.fun.\*
- 17: em.hum.\*
- 18: em.in.\*
- 19: em.mu.\*
- 20: em.om.\*
- 21: em.or.\*
- 22: em.ov.\*
- 23: em.pat.\*
- 24: em.ph.\*
- 25: em.pl.\*
- 26: em.ro.\*
- 27: em.sts.\*
- 28: em.un.\*
- 29: em.vi.\*
- 30: em.htg.hum.\*
- 31: em.htg.inv.\*
- 32: em.htg.other.\*
- 33: em.htg.mus.\*
- 34: em.htg.pin.\*
- 35: em.htg.rod.\*
- 36: em.htg.man.\*
- 37: em.htg.vrt.\*
- 38: em.sy.\*
- 39: em.htgo.hum.\*
- 40: em.htgo.mus.\*
- 41: em.htgo.other.\*

Pred. No. is the number of results predicted by chance to have a

Result No. Score Match Length DB ID Description

1	3025	100.0	3025	6	BD269559	Novel CAR
2	3025	100.0	3025	6	AR269111	Sequence
3	3025	100.0	3025	6	AR437545	Sequence
4	2980.4	98.5	3001	9	AF116826	Homo sapi
5	2980	82.9	2980	9	AY303691	Homo sapi
6	2508	82.9	2508	6	AX056437	Sequence
7	2505	82.8	2505	6	BD269560	Novel CAR
8	2505	82.8	2505	6	AR269112	Sequence
9	2505	82.8	2505	6	AR437546	Sequence
10	2010	66.4	3026	6	BD269561	Novel CAR
11	2010	66.4	3026	6	AR269113	Sequence
12	2010	66.4	3026	6	AR437547	Sequence
13	1991.4	65.8	2982	10	AY303692	Novel CAR
14	1892.2	62.6	2505	6	BD269562	Novel CAR
15	1892.2	62.6	2505	6	AR269114	Sequence
16	1892.2	62.6	2505	6	AR437548	Sequence
17	1730.6	57.2	7807	9	HSN807040	Homo sapi
18	524	17.3	75492	9	EX470253	Human DNA
19	524	17.3	164132	2	AC019284	Homo sapi
20	524	17.3	165947	2	AC015897	Homo sapi
21	466.4	15.4	668	6	AX525654	Sequence
22	199.4	6.6	150272	2	AC073887	Homo sapi
23	199.4	6.6	179582	9	AC093158	Homo sapi
24	196.2	6.5	66066	2	AC026566	Homo sapi
25	177.6	5.9	112358	9	AC105271	Homo sapi
26	166.6	5.5	197553	10	AC124174	Mus muscu
27	165	5.5	269035	2	AC126082	Rattus no
28	165	5.5	276946	2	AC115396	Rattus no
29	153.4	5.1	150272	2	AC073887	Homo sapi
30	153	5.1	146018	2	AC018404	Homo sapi
31	149.4	4.9	66066	2	AC026566	Homo sapi
32	141.2	4.7	178057	10	AC125097	Mus muscu
33	138.2	4.6	258378	2	AC105566	Rattus no
34	110.4	3.6	78331	9	AC098692	Homo sapi
35	107.2	3.5	60037	2	AC100314	Mus muscu
36	107.2	3.5	254819	2	AC144762	Mus muscu
37	106.4	3.5	1672	8	AY096470	Arabidops
38	106.4	3.5	2122	8	AY070086	Arabidops
39	101.4	3.4	28196	9	AC119672	Homo sapi
40	100.2	3.3	1704	6	AX652967	Sequence
41	100.2	3.3	2263	8	AK072690	Oryza sat
42	99.4	3.3	1748	8	AY128938	Arabidops
43	99.4	3.3	2104	8	AY093017	Arabidops
44	98	3.2	1662	6	AX506319	Sequence
45	94.8	3.1	1961	8	AK073747	Oryza sat

ALIGNMENTS

RESULT 1  
BD269559 3025 bp DNA linear PAT 17-JUL-2003  
LOCUS BD269559  
DEFINITION Novel CAR protein and nucleic acid molecules and uses therefor.  
ACCESSION BD269559  
VERSION BD269559.1 GI:33079327  
KEYWORDS JP 2002535962-A/1.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Bukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 3025)  
AUTHORS Raju, J.  
TITLE Novel CAR protein and nucleic acid molecules and uses therefor  
JOURNAL Patent: JP 2002535962-A 1 29-OCT-2002;



```
Db 1741 TTATTGAGTAGATGTTGCGCAAGGATGGAGTACCTTACACACCTGCACAGCCAAATTA 1800
Qy 1801 TACATCGTGAAGTGAACAGTCAACAATTTCTCTATGAGGATGGGATGCTGGTGG 1860
Db 1801 TACATCGTGAAGTGAACAGTCAACAATTTCTCTATGAGGATGGGATGCTGGTGG 1860
Qy 1861 CAGATTTTGGAGAAATCAAGATTTCTACAGTCTCTGGATGAAGACAAACATGACAAAC 1920
Db 1861 CAGATTTTGGAGAAATCAAGATTTCTACAGTCTCTGGATGAAGACAAACATGACAAAC 1920
Qy 1921 CTGGGAACCTCCGTTGGATGGCTTCCAGAGTGTTCACGAGTGCATCGGTACACCATCA 1980
Db 1921 CTGGGAACCTCCGTTGGATGGCTTCCAGAGTGTTCACGAGTGCATCGGTACACCATCA 1980
Qy 1981 AAGCAGATGTCTTCAGCTATGCTCTGTCTGTGGGAAATTTCTCACTGGCGAAATTTCCAT 2040
Db 1981 AAGCAGATGTCTTCAGCTATGCTCTGTCTGTGGGAAATTTCTCACTGGCGAAATTTCCAT 2040
Qy 2041 TCGCTCATCTCAAGCAGCGGCTGGCGGAGCAGACATGCGCTTACCACACATCAGACCTC 2100
Db 2041 TCGCTCATCTCAAGCAGCGGCTGGCGGAGCAGACATGCGCTTACCACACATCAGACCTC 2100
Qy 2101 CCATTGGCTATTCCATTCCCAAGCCCATATCATCTCTGCTGATAGAGGGTGGAAACGAT 2160
Db 2101 CCATTGGCTATTCCATTCCCAAGCCCATATCATCTCTGCTGATAGAGGGTGGAAACGAT 2160
Qy 2161 GTCTTGAAGAACACCGAATTTCTGAAAGTGTGTCATGAAGTTAGAGAGTGTCTCTGCA 2220
Db 2161 GTCTTGAAGAACACCGAATTTCTGAAAGTGTGTCATGAAGTTAGAGAGTGTCTCTGCA 2220
Qy 2221 ACATTGAGCTGATGTCCTCGCATCAAGTAAACAGCAGTGGTCTCTCACCTTCTCTT 2280
Db 2221 ACATTGAGCTGATGTCCTCGCATCAAGTAAACAGCAGTGGTCTCTCACCTTCTCTT 2280
Qy 2281 CTCTCATGTCCTTGGTGAACCGGGGAGGACCTTGGCGGAGTCAATGCGCAGCATTAAGAA 2340
Db 2281 CTCTCATGTCCTTGGTGAACCGGGGAGGACCTTGGCGGAGTCAATGCGCAGCATTAAGAA 2340
Qy 2341 GTCGTTTCGAATTTGGAATATGCTCTAAATGCAAGTCTTATGCTGCTTGTCCCAAGTG 2400
Db 2341 GTCGTTTCGAATTTGGAATATGCTCTAAATGCAAGTCTTATGCTGCTTGTCCCAAGTG 2400
Qy 2401 CTGACAAATATTCCTCTCAAGTCTGTCTTTGGAGGAGTGAAGAGTCTTCAATACA 2460
Db 2401 CTGACAAATATTCCTCTCAAGTCTGTCTTTGGAGGAGTGAAGAGTCTTCAATACA 2460
Qy 2461 CACCCATTGACAAATATGCTATGATCCGATCCCATGAGCTCAATGCAATTTCAATCTT 2520
Db 2461 CACCCATTGACAAATATGCTATGATCCGATCCCATGAGCTCAATGCAATTTCAATCTT 2520
Qy 2521 GCGGAATAGTACAGCTTTGAGGACAGCAGTGCAGCAGATTCGGGTATACCTAAGGAG 2580
Db 2521 GCGGAATAGTACAGCTTTGAGGACAGCAGTGCAGCAGATTCGGGTATACCTAAGGAG 2580
Qy 2581 AGTTTTTTTCCCGAACTGACAGCAACGATTTCCAAACCGGAGTGGCTTCCAACTATA 2640
Db 2581 AGTTTTTTTCCCGAACTGACAGCAACGATTTCCAAACCGGAGTGGCTTCCAACTATA 2640
Qy 2641 ACATTTTACTCTCAAGGCTCTCTAAATTTGGCTTTGTTTACTTGCTCTTAAATTC 2700
Db 2641 ACATTTTACTCTCAAGGCTCTCTAAATTTGGCTTTGTTTACTTGCTCTTAAATTC 2700
Qy 2701 CCACATATTAGCAGGCTTTGGATTTGCTTAAGGAATAATGCAAAAGAACCAAGACA 2760
Db 2701 CCACATATTAGCAGGCTTTGGATTTGCTTAAGGAATAATGCAAAAGAACCAAGACA 2760
Qy 2761 GAATGATATGAAGATTTGTTTAAATTTGTAATTTAAATTTAAATTTAGATCGTTACT 2820
Db 2761 GAATGATATGAAGATTTGTTTAAATTTGTAATTTAAATTTAAATTTAGATCGTTACT 2820
Qy 2821 TGGAAATGGAGCCTTACGCTGCTGGTGACAGATAAATATGTTTCTCGTGGCTGAATT 2880
```

```
Db 2821 TGGAAATGGAGCCTAAGTCTGTGTGGACAGATAAATATGTTTCTCGGCTGAATT 2880
Qy 2881 ATGTAGACTGTGTTTGACAGCTATGGGTTTATTTCTTAGAACATTTGTTCTTTT 2940
Db 2881 ATGTAGACTGTGTTTGACAGCTATGGGTTTATTTCTTAGAACATTTGTTCTTTT 2940
Qy 2941 CTCAATTATGTTACTTCTAGTGTTCACCTCTGTGATTAAGATTTCTTGTGTAATAGAAA 3000
Db 2941 CTCAATTATGTTACTTCTAGTGTTCACCTCTGTGATTAAGATTTCTTGTGTAATAGAAA 3000
Qy 3001 AAAAAAAGGCGGCGC 3025
Db 3001 AAAAAAAGGCGGCGC 3025

RESULT 2
AR269111 LOCUS 3025 bp DNA linear PAT 10-APR-2003
DEFINITION Sequence 1 from patent US 6500654.
ACCESSION AR269111
VERSION AR269111.1 GI:29699954
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 3025)
AUTHORS Raju,J.
TITLE CARP protein and nucleic acid molecules and uses therefor
JOURNAL Patent: US 6500654-A 1 31-DEC-2002;
FEATURES
source 1. .3025
/mol_type="genomic DNA"
ORIGIN

Query Match 100.0%; Score 3025; DB 6; Length 3025;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3025; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTGAGCCCGGCTCGCGCCCTGGAGAGGAGAACTTATAATAAATGGAAATTATA 60
Db 1 GTGAGCCCGGCTCGCGCCCTGGAGAGGAGAACTTATAATAAATGGAAATTATA 60
Qy 61 AATCTAGACCAACCCAAACTTGTACTGATGAATGGAAGAAAAGTCAATCATATG 120
Db 61 AATCTAGACCAACCCAAACTTGTACTGATGAATGGAAGAAAAGTCAATCATATG 120
Qy 121 TTATCAATAGAAAGATTAGAGATGACCTGCAGATCAAGGAAAAGACTGACAGAAC 180
Db 121 TTATCAATAGAAAGATTAGAGATGACCTGCAGATCAAGGAAAAGACTGACAGAAC 180
Qy 181 TAAGGAATATTTGGCTCTGTGAAGCCTTCAGTAAAGTCAATTTAAATACCCGACTG 240
Db 181 TAAGGAATATTTGGCTCTGTGAAGCCTTCAGTAAAGTCAATTTAAATACCCGACTG 240
Qy 241 AAAATGGGCTGCTCTCACTTTCAATTTATGTTGATTTGGAGGCAAGAAATCATAATTC 300
Db 241 AAAATGGGCTGCTCTCACTTTCAATTTATGTTGATTTGGAGGCAAGAAATCATAATTC 300
Qy 301 GAACTCTTATGTTGAAGGCTCCGCCCATCTCGACTGACAGAAATGGATTACAGCT 360
Db 301 GAACTCTTATGTTGAAGGCTCCGCCCATCTCGACTGACAGAAATGGATTACAGCT 360
Qy 361 TGCATTAGCAGTTTACAAGGATAATGCAAGATTGATCACTTCTCTGCTTCAAGTGAG 420
Db 361 TGCATTAGCAGTTTACAAGGATAATGCAAGATTGATCACTTCTCTGCTTCAAGTGAG 420
Qy 421 CTGATATACAGAGTTGGATGAGGTGGCTCAGTCCCTCCATTTGCTACAAATAGCTG 480
Db 421 CTGATATACAGAGTTGGATGAGGTGGCTCAGTCCCTCCATTTGCTACAAATAGCTG 480
Qy 481 GCCACCTAGAGGCTGCTGATGCTGCTTGCACATGGAGCTAATGTCATATTTCAAGATG 540
```

Db 481 GCCACCTAGAGGCTGCTGATGTGCTGTTGTAACATGGAGCTAATGTCAATATTTCAAGATG 540  
Qy 541 CAGTTTTTTTCACTCCATTGCATATTTGACGCTACTATGGACATGAACAGTAACTCGCC 600  
Db 541 CAGTTTTTTTCACTCCATTGCATATTTGACGCTACTATGGACATGAACAGTAACTCGCC 600  
Qy 601 TTTCTTTTGAATTTGGTCTGCTGATGTAATGTAAGTGGTGAAGTTGGAGTAGACCCCTCC 660  
Db 601 TTTCTTTTGAATTTGGTCTGCTGATGTAATGTAAGTGGTGAAGTTGGAGTAGACCCCTCC 660  
Qy 661 ACCTAGCATCTGCAAAAGGATTCCTTGAATATTTGAAAACCTCTTGATGGAAAGAGGCAGCA 720  
Db 661 ACCTAGCATCTGCAAAAGGATTCCTTGAATATTTGAAAACCTCTTGATGGAAAGAGGCAGCA 720  
Qy 721 AAGCAGATGTGAATGCTCAAGATAATGAAGACCAATGTCCTCACTTCTGTTCTCGAT 780  
Db 721 AAGCAGATGTGAATGCTCAAGATAATGAAGACCAATGTCCTCACTTCTGTTCTCGAT 780  
Qy 781 TTGGACACCATGATATAGTAAAGTATCTGCTGCAAGTGAATTTGGAAGTTCAACCTCATG 840  
Db 781 TTGGACACCATGATATAGTAAAGTATCTGCTGCAAGTGAATTTGGAAGTTCAACCTCATG 840  
Qy 841 TTGTTAAATCTATGGAGATACCCCTTACACCTGGCATGCTACAAATGGCAAAATTTGAAG 900  
Db 841 TTGTTAAATCTATGGAGATACCCCTTACACCTGGCATGCTACAAATGGCAAAATTTGAAG 900  
Qy 901 TTGCCAAGGAATCATCCAAATATCAGGAACAGAAAGTCTGACTAAGGAACATCTTCA 960  
Db 901 TTGCCAAGGAATCATCCAAATATCAGGAACAGAAAGTCTGACTAAGGAACATCTTCA 960  
Qy 961 GTGAAACAGCTTTTCATAGTCTTTGTACCTATGGCAAGACATTTGACCTAGTCAAAATTTT 1020  
Db 961 GTGAAACAGCTTTTCATAGTCTTTGTACCTATGGCAAGACATTTGACCTAGTCAAAATTTT 1020  
Qy 1021 TTTCTGATCAGAAATGTCAATAACATCAACCAAGGAAGGATGGGCACATGGATTAC 1080  
Db 1021 TTTCTGATCAGAAATGTCAATAACATCAACCAAGGAAGGATGGGCACATGGATTAC 1080  
Qy 1081 ACTCTGCTGTACACAGGTGATTCGCTGTTTCACTTCTTACTGGATTAATGGAGCTG 1140  
Db 1081 ACTCTGCTGTACACAGGTGATTCGCTGTTTCACTTCTTACTGGATTAATGGAGCTG 1140  
Qy 1141 ATATGAATCTAGTGGCTTGTATCCAGCAGGCTTAGTGTGAAAAGATGAGCAGACAT 1200  
Db 1141 ATATGAATCTAGTGGCTTGTATCCAGCAGGCTTAGTGTGAAAAGATGAGCAGACAT 1200  
Qy 1201 GTTTGATGTGGCTTTATGAAAAGGCGATGATGCTTGTACACTTCTTCAAGCATTTATA 1260  
Db 1201 GTTTGATGTGGCTTTATGAAAAGGCGATGATGCTTGTACACTTCTTCAAGCATTTATA 1260  
Qy 1261 AGAGACCAAGATGAAATTTGCTTGAATGAATTTCTCAGCTGGAGGAGATGGCTCT 1320  
Db 1261 AGAGACCAAGATGAAATTTGCTTGAATGAATTTCTCAGCTGGAGGAGATGGCTCT 1320  
Qy 1321 ATGTGTCTGTTTCCATCACCCTTGGGGAAGATTAAGCATGACAAAAGAGGCGAGATA 1380  
Db 1321 ATGTGTCTGTTTCCATCACCCTTGGGGAAGATTAAGCATGACAAAAGAGAGGCGAGATA 1380  
Qy 1381 TTTCTCTCTAAGAGCTGGATTTGCTTTCACATTTCCATTTCCATTTCCATTTTGAATTTGAT 1440  
Db 1381 TTTCTCTCTAAGAGCTGGATTTGCTTTCACATTTCCATTTCCATTTTGAATTTGAT 1440  
Qy 1441 TCCATGATTAATTTGGCTCAGGTTCTTTTGGGAAAGTATATAAGGACGATGAGAAATA 1500  
Db 1441 TCCATGATTAATTTGGCTCAGGTTCTTTTGGGAAAGTATATAAGGACGATGAGAAATA 1500  
Qy 1501 AAATAGTGGCTATATAAGCTTATCGAGCCTTACTCTCTCAAGTCAAGTGTGGATA 1560  
Db 1501 AAATAGTGGCTATATAAGCTTATCGAGCCTTACTCTCTCAAGTCAAGTGTGGATA 1560  
Qy 1561 TGTTTTGGCAGAGGTTGCTATCTCTGCGAGCTCAATCATCTCCCTCGTAATTCAGTTTG 1620  
Db 1561 TGTTTTGGCAGAGGTTGCTATCTCTGCGAGCTCAATCATCTCCCTCGTAATTCAGTTTG 1620

Qy 1621 TGGGTGCTTCTGCTGAATGATCCAGCGAGTTTGGCCATTTGCTCACTCAATACATATCAGGGG 1680  
Db 1621 TGGGTGCTTCTGCTGAATGATCCAGCGAGTTTGGCCATTTGCTCACTCAATACATATCAGGGG 1680  
Qy 1681 GTTCTCTGTTCTCCCTCTTTCATGACGAAGAGGATTTCTGATTGTCAGTCTAAATTA 1740  
Db 1681 GTTCTCTGTTCTCCCTCTTTCATGACGAAGAGGATTTCTGATTGTCAGTCTAAATTA 1740  
Qy 1741 TTATTTGAGTGTGTTGTCACAAAGGATGAGTACCTTCAACACCTGACACAGCCAAATTA 1800  
Db 1741 TTATTTGAGTGTGTTGTCACAAAGGATGAGTACCTTCAACACCTGACACAGCCAAATTA 1800  
Qy 1801 TACATCTGATCTTGAACAGTCAAAATTTCTTCTCTATGAGGATGGGATGCTGTGGTGG 1860  
Db 1801 TACATCTGATCTTGAACAGTCAAAATTTCTTCTCTATGAGGATGGGATGCTGTGGTGG 1860  
Qy 1861 CAGATTTTGGAGAAATCAAGATTTTCTACAGTCTCTGGATGAAGCAACATGACAAACAAAC 1920  
Db 1861 CAGATTTTGGAGAAATCAAGATTTTCTACAGTCTCTGGATGAAGCAACATGACAAACAAAC 1920  
Qy 1921 CTGGGAACCTCCGTTGGATGGCTCTGAGGTGTTACGCGAGTGCATCTGGTACACCATCA 1980  
Db 1921 CTGGGAACCTCCGTTGGATGGCTCTGAGGTGTTACGCGAGTGCATCTGGTACACCATCA 1980  
Qy 1981 AAGCAGATGTCTTCTGATGCTCTGCTGTGTTGGGAAATTTCTCACTGGCGAAATTTCCAT 2040  
Db 1981 AAGCAGATGTCTTCTGATGCTCTGCTGTGTTGGGAAATTTCTCACTGGCGAAATTTCCAT 2040  
Qy 2041 TCGTCTATCTCAAGCAGCGGCTGCGGAGCAGACATGAGCTTTACCAACATCAGACCTC 2100  
Db 2041 TCGTCTATCTCAAGCAGCGGCTGCGGAGCAGACATGAGCTTTACCAACATCAGACCTC 2100  
Qy 2101 CCATTTGGCTATTTCCATTTCCAGCCCATATCTCTGCTGATAGAGGTTGGAACGAT 2160  
Db 2101 CCATTTGGCTATTTCCATTTCCAGCCCATATCTCTGCTGATAGAGGTTGGAACGAT 2160  
Qy 2161 GTCTGTAAGGAAGACCCGAAATTTCTGAAAGTTGTCTATGAAGTTAGAAGAGTGTCTCTCA 2220  
Db 2161 GTCTGTAAGGAAGACCCGAAATTTCTGAAAGTTGTCTATGAAGTTAGAAGAGTGTCTCTCA 2220  
Qy 2221 ACATTTGAGCTGATGCTCTCTGATCAAGTAAACAGAGTGGTCTCTCTCACCTCTCTT 2280  
Db 2221 ACATTTGAGCTGATGCTCTCTGATCAAGTAAACAGAGTGGTCTCTCTCACCTCTCTT 2280  
Qy 2281 CTTCTGATTTCCCTGCTGTAACCCGCGAGGACCTGGCGGAGTCACTGTCGAGCATTAAGAA 2340  
Db 2281 CTTCTGATTTCCCTGCTGTAACCCGCGAGGACCTGGCGGAGTCACTGTCGAGCATTAAGAA 2340  
Qy 2341 GTCGTTTTCGAATTTGGAATATGCTCTAAATGCAAGGTCTTATGCTCTTTGTCGCGAAGTG 2400  
Db 2341 GTCGTTTTCGAATTTGGAATATGCTCTAAATGCAAGGTCTTATGCTCTTTGTCGCGAAGTG 2400  
Qy 2401 CTGGAACAATATTTCTCTCAAGGTCGTCTTTGGAGAGATGAAAAGAGTCTTCAATACA 2460  
Db 2401 CTGGAACAATATTTCTCTCAAGGTCGTCTTTGGAGAGATGAAAAGAGTCTTCAATACA 2460  
Qy 2461 CACCCATTTGAACAATATGCTATGATCCGATCCCATGAGCTCAATGATTTTCTTCTT 2520  
Db 2461 CACCCATTTGAACAATATGCTATGATCCGATCCCATGAGCTCAATGATTTTCTTCTT 2520  
Qy 2521 GCCGAATATGATGAGCTTTGAGGACAGCAGCTGACAGATTTGGCGGTATACCTTAAGGAG 2580  
Db 2521 GCCGAATATGATGAGCTTTGAGGACAGCAGCTGACAGATTTGGCGGTATACCTTAAGGAG 2580  
Qy 2581 AGTTTTTTTCCCGAACTGACAGCAACGATTTCCAAACCGGCAAGCTGCTTCCAACTATA 2640  
Db 2581 AGTTTTTTTCCCGAACTGACAGCAACGATTTCCAAACCGGCAAGCTGCTTCCAACTATA 2640  
Qy 2641 ACATTTTACTCTCAAGAGTCTCTTAAATTTGGGCTTGTGTTTACTGTTCTTAAATTC 2700  
Db 2641 ACATTTTACTCTCAAGAGTCTCTTAAATTTGGGCTTGTGTTTACTGTTCTTAAATTC 2700

Qy	2701	CCCACTATTAGCAGCGCTTTGGATTGTGCTTACAGGAATAATATGCAAAAGAACCAAGACA	2760			
Db	2701	CCCACTATTAGCAGCGCTTTGGATTGTGCTTACAGGAATAATATGCAAAAGAACCAAGACA	2760			
Qy	2761	GAATGTATATGAAGAATTTGTTTTTAAATTTTGTAAATTTAAAAAAAATTTAGATCGTTACT	2820			
Db	2761	GAATGTATATGAAGATTTGTTTTTAAATTTTGTAAATTTAAAAAAAATTTAGATCGTTACT	2820			
Qy	2821	TGGAATATGGAGCCTAAGTCTGTGTGGACAGATAATAATATATGTTTTCTCGGCTGAATT	2880			
Db	2821	TGGAATATGGAGCCTAAGTCTGTGTGGACAGATAATAATATATGTTTTCTCGGCTGAATT	2880			
Qy	2881	ATGTAGACTTGTGTTTGACACGATATGGTTTTATTTCTTAGACACATGTTTCATTTTCTTTT	2940			
Db	2881	ATGTAGACTTGTGTTTGACACGATATGGTTTTATTTCTTAGACACATGTTTCATTTTCTTTT	2940			
Qy	2941	CTCATTTAGTTACTTCTAGTGTTCACCTCTGTGATTAAGAATCTTTGGTGAAATAGAAA	3000			
Db	2941	CTCATTTAGTTACTTCTAGTGTTCACCTCTGTGATTAAGAATCTTTGGTGAAATAGAAA	3000			
Qy	3001	AAAAAAAAGGCGGCCGC	3025			
Db	3001	AAAAAAAAGGCGGCCGC	3025			
RESULT 3						
AR437545						
LOCUS	AR437545	3025 bp	DNA linear PAT 18-DEC-2003			
DEFINITION	Sequence 1 from patent US 6660490.					
ACCESSION	AR437545					
VERSION	AR437545.1 GI:40202640					
KEYWORDS	Unknown.					
SOURCE	Unknown.					
ORGANISM	Unclassified.					
REFERENCE	1 (bases 1 to 3025)					
AUTHORS	Raju, J.					
TITLE	CARP protein and nucleic acid molecules and uses therefor					
JOURNAL	Patent: US 6660490-A 1 09-DEC-2003;					
FEATURES	Location/Qualifiers					
source	1..3025					
/organism="unknown"						
/mol_type="genomic DNA"						
ORIGIN						
Query Match 100.0%; Score 3025; DB 6; Length 3025;						
Best Local Similarity 100.0%; Pred. No. 0;						
Matches 3025; Conservative 0; Mismatches 0; Indels 0; Gaps 0;						
Qy	1	GTGACCCACGCGTCCGCCCTCGAGAAAGAAAGAACTTATAATAATGGAAATTATA	60			
Db	1	GTGACCCACGCGTCCGCCCTCGAGAAAGAAAGAACTTATAATAATGGAAATTATA	60			
Qy	61	AATCTAGACCAACCAAACTTGACTGATGGAAGAAAAGTCAAGTGAATCATATG	120			
Db	61	AATCTAGACCAACCAAACTTGACTGATGGAAGAAAAGTCAAGTGAATCATATG	120			
Qy	121	TTATCACAATAGAAAGATTAGAAGATCACTTCAGATCAAGGAAAAAGAACTGCAGAAC	180			
Db	121	TTATCACAATAGAAAGATTAGAAGATCACTTCAGATCAAGGAAAAAGAACTGCAGAAC	180			
Qy	181	TAAAGAAATATTTGGCTCTGATGAAGCCTTCAGTAAAGTCAATTTAAATTCGCGACTG	240			
Db	181	TAAAGAAATATTTGGCTCTGATGAAGCCTTCAGTAAAGTCAATTTAAATTCGCGACTG	240			
Qy	241	AAAAATGGGCTGCTCTACTTTCATTTATGTTGCAATTTGTGAGGCAAGAAATCACATATTC	300			
Db	241	AAAAATGGGCTGCTCTACTTTCATTTATGTTGCAATTTGTGAGGCAAGAAATCACATATTC	300			
Qy	301	GAACTCTTATGTTGAAAGGGCTCCGCCCATCTCGACTGACAAGAAATGGAATTCAGCGCT	360			
Db	301	GAACTCTTATGTTGAAAGGGCTCCGCCCATCTCGACTGACAAGAAATGGAATTCAGCGCT	360			

```
Db 1441 TCCATGAGATTTATGGCTCAGGTTCTTTTGGAAAGTATATAAAGGACGATGAGAAATA 1500
QY 1501 AAATAGTGGCTATAAAAGCTTATCGAGCAATACCTACTGCTCCAAAGTCAGATGTGGATA 1560
Db 1501 AAATAGTGGCTATAAAAGCTTATCGAGCAATACCTACTGCTCCAAAGTCAGATGTGGATA 1560
QY 1561 TGTCTTGGCGAGAGTGTCATCTCTGCGAGCTCAATCATCCCTGGTAAATTCAGTTTG 1620
Db 1561 TGTCTTGGCGAGAGTGTCATCTCTGCGAGCTCAATCATCCCTGGTAAATTCAGTTTG 1620
QY 1621 TGGGTCCTTGTCTGAATGATCCAGCCAGTTTGGCAATGTGCACTCAATATCAGGGG 1680
Db 1621 TGGGTCCTTGTCTGAATGATCCAGCCAGTTTGGCAATGTGCACTCAATATCAGGGG 1680
QY 1681 GTTCTGTGTTCTCCCTCTCATGAGCAAGAGAGATCTTGATTTGCACTCAATATA 1740
Db 1681 GTTCTGTGTTCTCCCTCTCATGAGCAAGAGAGATCTTGATTTGCACTCAATATA 1740
QY 1741 TTATTGCACTAGATGTTGCCAAAGCATGGAGTACCTTCAACCTGACACAGCAATTA 1800
Db 1741 TTATTGCACTAGATGTTGCCAAAGCATGGAGTACCTTCAACCTGACACAGCAATTA 1800
QY 1801 TACATCGTGAATGTTGAACAGTCAATATCTTCTATGAGGATGGGCATGCTGTGTGG 1860
Db 1801 TACATCGTGAATGTTGAACAGTCAATATCTTCTATGAGGATGGGCATGCTGTGTGG 1860
QY 1861 CAGATTTGGAGATCAAGATTTCTACAGTCTCTGGATGAAGCAACATCACAAAAACA 1920
Db 1861 CAGATTTGGAGATCAAGATTTCTACAGTCTCTGGATGAAGCAACATCACAAAAACA 1920
QY 1921 CTGGGAACCTCGTTGGATGCTCCTGAGGTGTTTCAAGCATGCTGCTGATCAGCATCA 1980
Db 1921 CTGGGAACCTCGTTGGATGCTCCTGAGGTGTTTCAAGCATGCTGCTGATCAGCATCA 1980
QY 1981 AAGCAGATGCTTTCAGTATGCTCTGTTGTTGGGAAATTTCTACCTGGCGAAATTCCT 2040
Db 1981 AAGCAGATGCTTTCAGTATGCTCTGTTGTTGGGAAATTTCTACCTGGCGAAATTCCT 2040
QY 2041 TCGCTCATCTCAAGCCAGCGCTCGGCAGCAGACATGGCTTACCACCATCAGACCTC 2100
Db 2041 TCGCTCATCTCAAGCCAGCGCTCGGCAGCAGACATGGCTTACCACCATCAGACCTC 2100
QY 2101 CCAATTGGCTATTCATCCCAAGCCCATATCATCTCTGATACGAGGTGGAACGCAT 2160
Db 2101 CCAATTGGCTATTCATCCCAAGCCCATATCATCTCTGATACGAGGTGGAACGCAT 2160
QY 2161 GTCCCTGAAGGAAGCCCGAATTTCTGAAGTTGTCATGAAGTTAGAGAGTGCTCTGCA 2220
Db 2161 GTCCCTGAAGGAAGCCCGAATTTCTGAAGTTGTCATGAAGTTAGAGAGTGCTCTGCA 2220
QY 2221 ACATTGAGCTGATGCTCCTGCATCAAGTAAACAGCAGTGGGTCTCTCTCACCTTCCT 2280
Db 2221 ACATTGAGCTGATGCTCCTGCATCAAGTAAACAGCAGTGGGTCTCTCTCACCTTCCT 2280
QY 2281 CTTCTGATGCTGGTGAACCGGGAGCCTGCGGAGTCATGTGGCAGATTAAAGA 2340
Db 2281 CTTCTGATGCTGGTGAACCGGGAGCCTGCGGAGTCATGTGGCAGATTAAAGA 2340
QY 2341 GTCGTTTCCGAATTGGAATATGCTCTAAATGCAAGTCTTATGCTGCTTTGTCCTCAAG 2400
Db 2341 GTCGTTTCCGAATTGGAATATGCTCTAAATGCAAGTCTTATGCTGCTTTGTCCTCAAG 2400
QY 2401 CTGGCAATATCTCTCAAGTCTGCTTTGGAGGAGTGAAGAGTCTTCAATCA 2460
Db 2401 CTGGCAATATCTCTCAAGTCTGCTTTGGAGGAGTGAAGAGTCTTCAATCA 2460
QY 2461 CACCCATTGACAAATATGGCTATGATCCGATCCCATGAGCTCAATGCAATTTTCATCT 2520
Db 2461 CACCCATTGACAAATATGGCTATGATCCGATCCCATGAGCTCAATGCAATTTTCATCT 2520
QY 2521 GCCGAAATAGTAGAGCTTTGAGGACAGCAGCTGACAGCATTTGGGTATACCTAAGGAG 2580
```

```
Db 2521 GCCGAAATAGTAGAGCTTTGAGGACAGCAGCTGACAGCATTCGGCGTATACCTAAGGAG 2580
QY 2581 AGTTTTTTTCCCCGAACTGACAGCAACGATTCACACCGCAAGCTGGCTTCCAACTATA 2640
Db 2581 AGTTTTTTTCCCCGAACTGACAGCAACGATTCACACCGCAAGCTGGCTTCCAACTATA 2640
QY 2641 ACATTTTACTCTCAAGGCTCTCCTTAAATTTGGGCTTGTCTTACTTGTCTCTTAAATTC 2700
Db 2641 ACATTTTACTCTCAAGGCTCTCCTTAAATTTGGGCTTGTCTTACTTGTCTCTTAAATTC 2700
QY 2701 CCCACTATTAGCAGGCTTTGGAATTTGTGCTCAAGGAATAATATGCAAAAGAACCAAGACA 2760
Db 2701 CCCACTATTAGCAGGCTTTGGAATTTGTGCTCAAGGAATAATATGCAAAAGAACCAAGACA 2760
QY 2761 GAATGTATATGAAGATTTGTTTTAAATTTGTAAATTTAAATAAATAAATTTAGATCGTTACT 2820
Db 2761 GAATGTATATGAAGATTTGTTTTAAATTTGTAAATTTAAATAAATAAATTTAGATCGTTACT 2820
QY 2821 TGGAAATGAGCCCTAAGTCTGTGTGAGCAGATATAATATTTCTTCTGGCTGGAAT 2880
Db 2821 TGGAAATGAGCCCTAAGTCTGTGTGAGCAGATATAATATTTCTTCTGGCTGGAAT 2880
QY 2881 ATGTAGACTTTGTGTTGACAGCTATGGGTTTATTTCTTGAACATTTGTCATTTCTTTT 2940
Db 2881 ATGTAGACTTTGTGTTGACAGCTATGGGTTTATTTCTTGAACATTTGTCATTTCTTTT 2940
QY 2941 CTCATTATGTTACTTCTAGTGTTCACCTCTGTGATTAAAGATTTCTTGGTGAATAGAAA 3000
Db 2941 CTCATTATGTTACTTCTAGTGTTCACCTCTGTGATTAAAGATTTCTTGGTGAATAGAAA 3000
QY 3001 AAAAAAATAAAGAGCGCGCGC 3025
Db 3001 AAAAAAATAAAGAGCGCGCGC 3025

RESULT 4
AF116826 Homo sapiens clone HH498 putative protein-tyrosine kinase mRNA, linear PRI 11-MAY-1999
LOCUS DEFINITION complete cds.
ACCESSION AF116826
VERSION AF116826.1 GI:4768828
KEYWORDS FLI CDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 3001)
AUTHORS Wei,Y.J., Ding,J.F., Xiong,H., Zhou,Y., Hui,R.T. and Liew,C.C.
TITLE Direct Submission
JOURNAL Submitted (25-DEC-1998) Molecular Medicine Center, Cardiovascular
Institute and Fu Wai Hospital, 167 Beilishi Road, West District,
Beijing 100037, China
FEATURES
Location/Qualifiers
1..3001
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="heart"
/dev_stage="adult"
52..2559
/codon_start=1
/product="putative protein-tyrosine kinase"
/protein_id="AAD29632.1"
/db_xref="GI:4768828"
/translation="MGNYKSRFTQCTDEWKKKSVIETIERLEDLQIKKELTE
LNIFGSDAFKVNLYNRTENGSLHLCCCTCGGKSHIRTLMLKGRPLRTNGF
TALHLAVYKDNALITSLHSGADIQQVYGGTALHTATAGHLEADVLQHGANY
NQDAVFPTPLHIAAYVGEHQVTRLLKFGADVNSGEVDRPLHLAGAKFLNIAKL
LMEGSKADVNQDNEDHVPJLHFCRFGHHIDIVKYLQSDLEVOPHVNIYGDTPHL
ACYNGKFEVAKBIQISGESLTKENIPSETAFHSACTYKSIDLVKLLQDNVININ
HQGRDGHGLHSACTHGHRLVQLFDNGADNVLVACDPSRSSGKEQTCMLWAYEK
```

GHDAIVTLKXHKRQDELPCNEYSQPGDGSYVSPSLGKIKSMKXKADILLRA  
GLPSFHLQLSIEHEIIGSGFKYKGRNKAIVAKRYANTYCSKSDVMPGR  
EVSILCOLNHPVIOFVGACLNDSQFAIVTQYSGSLFSLHEKRIILDQSKLI  
AVDVARKMEYLNLTQPIIHRDLNHNILLYEDGHVAVDFESRFLQSLDENVTQ  
PGLNRMAPEVFTQCTRYTIKADVSALCIWEILTGEIPEAHLKPAAPAAADMAHYHI  
RPIGVSIKPISSILLIRGNWACPSRPEFSEVVMKLEECNLNKLMSPASSNSGSL  
SPSSSDCLVNRGGPRSHVAALRSRFELEVALNABSVAAALSQAGQYSSQGLSLEM  
KSLQTPIDIKTYGVSDPMSMHPHSCRNSSFEDESS"  
2978. .2983

polyA\_signal

ORIGIN

Query Match 98.5%; Score 2980.4; DB 9; Length 3001;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2981; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	16	CGGCCCTGAGAAAGGAACCTTATAATAAATGGGAAATATATAATCTAGACAACCC	75
DB	20	CTGCCCTGGAGAAAGGAACCTTATAATAAATGGGAAATATATAATCTAGACAACCC	79
QY	76	AAACITTTGTAAGTGAATGGAAGAAAAGTCAAGTGAATCATATGTTATCAATAGAAA	135
DB	80	AAACITTTGTAAGTGAATGGAAGAAAAGTCAAGTGAATCATATGTTATCAATAGAAA	139
QY	136	GATTAGAAGATGACCTGCAGATCAAGGAAAAAGAACTGACAGAACTAAGGAATATATTG	195
DB	140	GATTAGAAGATGACCTGCAGATCAAGGAAAAAGAACTGACAGAACTAAGGAATATATTG	199
QY	196	GCTCTGATGAGCCCTTCAGTAAAGTCAATTAATACCGCACTGAAATGGCTGCTC	255
DB	200	GCTCTGATGAGCCCTTCAGTAAAGTCAATTAATACCGCACTGAAATGGCTGCTC	259
QY	256	TACTTCATTTATGTTGTCATTTGGAGGCAAGAAATCAATATTCGAACTCTTATGTTGA	315
DB	260	TACTTCATTTATGTTGTCATTTGGAGGCAAGAAATCAATATTCGAACTCTTATGTTGA	319
QY	316	AAGGCTCCGCCCATCTCGACTCAAGAAATGGATTTACGCCCTTGCAATTAGCAGTTT	375
DB	320	AAGGCTCCGCCCATCTCGACTCAAGAAATGGATTTACGCCCTTGCAATTAGCAGTTT	379
QY	376	ACAAGATATGAGAAATGATCACTCTCTGCTTCCAGTGGAGCTGATATACAGCAGG	435
DB	380	ACAAGATATGAGAAATGATCACTCTCTGCTTCCAGTGGAGCTGATATACAGCAGG	439
QY	436	TTGGATACGGTGGCCCTCAGTCCCTCCATATTTGCTACAAATAGCTGGCCACCTAGAGGTG	495
DB	440	TTGGATACGGTGGCCCTCAGTCCCTCCATATTTGCTACAAATAGCTGGCCACCTAGAGGTG	499
QY	496	CTGATGTGCTTTGCAACATGAGCTAAATGTCAAATTCAGATGCAAGTCAAGTCTTTCATC	555
DB	500	CTGATGTGCTTTGCAACATGAGCTAAATGTCAAATTCAGATGCAAGTCAAGTCTTTCATC	559
QY	556	CATTGCATATTGACGGGTACTATGGACATGAACAGGTAACTCGCCCTCTTTTGAATTTG	615
DB	560	CATTGCATATTGACGGGTACTATGGACATGAACAGGTAACTCGCCCTCTTTTGAATTTG	619
QY	616	GTGCTGATGTAATATGTAAGTGGTAAGTGGAGATAGACCCCTCCACCTAGCATCTGCAA	675
DB	620	GTGCTGATGTAATATGTAAGTGGTAAGTGGAGATAGACCCCTCCACCTAGCATCTGCAA	679
QY	676	AAGGATTTGTAATATTGCAAACTCTTGAATGGAAGCAAGCAAGCAGATGTAATG	735
DB	680	AAGGATTTGTAATATTGCAAACTCTTGAATGGAAGCAAGCAAGCAGATGTAATG	739
QY	736	CTCAAGATAATGAAACCATGTCCCACTCCATTTCTGTTTCTCGATTTGGACACCATGATA	795
DB	740	CTCAAGATAATGAAACCATGTCCCACTCCATTTCTGTTTCTCGATTTGGACACCATGATA	799
QY	796	TAGTAAAGTATCTGCTCAAGTGAATTTGGAAGTTCAACTCATGTGTGTTATATCTATG	855
DB	800	TAGTAAAGTATCTGCTCAAGTGAATTTGGAAGTTCAACTCATGTGTGTTATATCTATG	859
QY	856	GAGATACCCCTTACACCTGGCATGTGTAATGGCAAAATTTGAAGTTGCCAAGGAATCA	915

DB	860	GAGATACCCCTTTACACCTGGCATCTACAATGGCAAAATTTGAAGTTGCCAAGGAATCA	919
QY	916	TCCAAATATCAGAAACAGAAAGTCTGACTAAGGAAAAACATCTTCTAGTGAACACGCTTTC	975
DB	920	TCCAAATATCAGAAACAGAAAGTCTGACTAAGGAAAAACATCTTCTAGTGAACACGCTTTC	979
QY	976	ATAGTGTCTTGTACCTATGGCAAGACATTTGACCTAGTCAAAATTTCTTCTTGTATCAGAATG	1035
DB	980	ATAGTGTCTTGTACCTATGGCAAGACATTTGACCTAGTCAAAATTTCTTCTTGTATCAGAATG	1039
QY	1036	TCATAAACATCAACCAAGGAAGGATGGGCACTCTGGAATTAACATCTCTGCTTGTCTTACC	1095
DB	1040	TCATAAACATCAACCAAGGAAGGATGGGCACTCTGGAATTAACATCTCTGCTTGTCTTACC	1099
QY	1096	AGGTCACATTCGCCCTGGTTCAGTTCTTACTGGATAATGGAGCTGATATGAATCTAGTGG	1155
DB	1100	AGGTCACATTCGCCCTGGTTCAGTTCTTACTGGATAATGGAGCTGATATGAATCTAGTGG	1159
QY	1156	CTTGTGATCCCAGCAGGCTTAGTGTGAAAAAGATGAGCAGACATGTTTGTATGTGGGCTT	1215
DB	1160	CTTGTGATCCCAGCAGGCTTAGTGTGAAAAAGATGAGCAGACATGTTTGTATGTGGGCTT	1219
QY	1216	ATGAAAAAGGGCATGATGCCATTGTCACTCTGGAAGCATTAAGAGACCAACAAGATG	1275
DB	1220	ATGAAAAAGGGCATGATGCCATTGTCACTCTGGAAGCATTAAGAGACCAACAAGATG	1279
QY	1276	AATTGCCCTGTAATGAATATCTCAGCCTGGAGGAGATGGCTCTCTATGTCTGTTCAT	1335
DB	1280	AATTGCCCTGTAATGAATATCTCAGCCTGGAGGAGATGGCTCTCTATGTCTGTTCAT	1339
QY	1336	CACCTTGGGGAGATTAAGACATGACAAAGAGAGAGGAGATATCTCTCTTAAGAG	1395
DB	1340	CACCTTGGGGAGATTAAGACATGACAAAGAGAGAGGAGATATCTCTCTTAAGAG	1399
QY	1396	CTGGATTGCTTCCATTTCCATCTCAGCTCTCAGAAATAGGTTCCATGAGATTATTG	1455
DB	1400	CTGGATTGCTTCCATTTCCATCTCAGCTCTCAGAAATAGGTTCCATGAGATTATTG	1459
QY	1456	GCTCAGGTTCTTTTGGGAAAGTATATAAGGACGATGAGAAATAAATAGTGGCTATAA	1515
DB	1460	GCTCAGGTTCTTTTGGGAAAGTATATAAGGACGATGAGAAATAAATAGTGGCTATAA	1519
QY	1516	AAGTTTATCGAGCAATACCTACTGCTCCAGTCTAGATGTCGATATGTTTGCAGAGG	1575
DB	1520	AAGTTTATCGAGCAATACCTACTGCTCCAGTCTAGATGTCGATATGTTTGCAGAGG	1579
QY	1576	TGTCCATTCTCTGCCAGCTCAATCATCCCTGCGTAATTCAGTTTGTGGGTGCTTCTTGA	1635
DB	1580	TGTCCATTCTCTGCCAGCTCAATCATCCCTGCGTAATTCAGTTTGTGGGTGCTTCTTGA	1639
QY	1636	ATGATCCCAGCCAGTTTGGCATTTGTCATCAATATATCAGGGGGTTCCTGTTCTCCC	1695
DB	1640	ATGATCCCAGCCAGTTTGGCATTTGTCATCAATATATCAGGGGGTTCCTGTTCTCCC	1699
QY	1696	TCCTTCATGAGCAGAGAGGATTTCTGATTTCAGTCTAAATTAATTAATTCAGTAGATG	1755
DB	1700	TCCTTCATGAGCAGAGAGGATTTCTGATTTCAGTCTAAATTAATTAATTCAGTAGATG	1759
QY	1756	TTCCAAAAGGATGAGTACCTTCAACCTGACACAGCAATTAATCATCTGTGACTTGA	1815
DB	1760	TTCCAAAAGGATGAGTACCTTCAACCTGACACAGCAATTAATCATCTGTGACTTGA	1819
QY	1816	ACAGTCACATATTTCTCTCTATGAGGATGGGATGCTGTGTGGCAGATTTTGGAGAT	1875
DB	1820	ACAGTCACATATTTCTCTCTATGAGGATGGGATGCTGTGTGGCAGATTTTGGAGAT	1879
QY	1876	CAAGATTTCTTACAGTCTCTGGATGAAGCAACATGACAAAACACCTGGGAACCTCCGTT	1935
DB	1880	CAAGATTTCTTACAGTCTCTGGATGAAGCAACATGACAAAACACCTGGGAACCTCCGTT	1939
QY	1936	GGATGGCTCTTACAGTCTCTGGATGACCTCGGTACACCATCAAGCAGATGCTTCA	1995
DB	1940	GGATGGCTCTTACAGTCTCTGGATGACCTCGGTACACCATCAAGCAGATGCTTCA	1999

[illegible]

Db 301 GGGCTCGGCCCATCTCGACTGACAAGAAATGGATTTACAGCCTTGCAATTTAGCAGTTTAC 360  
Qy 378 AAGGATAATGAGAAATGATCATCTTCTGTCTTACAGTGGAGCTGATATACAGCAGGTT 437  
Db 361 AAGGATAATGAGAAATGATCATCTTCTGTCTTACAGTGGAGCTGATATACAGCAGGTT 420  
Qy 438 GGATACGGTGGCTCACTGCGCTCCATATTGCTACAAATAGCTGGCCACCTAGAGCTGCT 497  
Db 421 GGATACGGTGGCTCACTGCGCTCCATATTGCTACAAATAGCTGGCCACCTAGAGCTGCT 480  
Qy 498 GATGCTGTTGCAACATGAGAGTAATGTCATATTTCAAGATGCAGTATTTTCTACTCCA 557  
Db 481 GATGCTGTTGCAACATGAGAGTAATGTCATATTTCAAGATGCAGTATTTTCTACTCCA 540  
Qy 558 TTGCATATTGACGCTACTATGACATGAACAGGTAACCTCGCTTCTTTTGAATTTGGT 617  
Db 541 TTGCATATTGACGCTACTATGACATGAACAGGTAACCTCGCTTCTTTTGAATTTGGT 600  
Qy 618 GCTGATGTAATGTAAGTGTGAAGTGGAGATAGACCCCTCCACCTAGCATCTGCAAAA 677  
Db 601 GCTGATGTAATGTAAGTGTGAAGTGGAGATAGACCCCTCCACCTAGCATCTGCAAAA 660  
Qy 678 GGATTTCTGTAATTTGCAAACTCTTGATGGAAGAGGAGCAAGACAGATGTGAATGCT 737  
Db 661 GGATTTCTGTAATTTGCAAACTCTTGATGGAAGAGGAGCAAGACAGATGTGAATGCT 720  
Qy 738 CAAGATAATGAAGACCATGCTCCACTCCATTTCTGTTCTGATTTGGACACCATGATATA 797  
Db 721 CAAGATAATGAAGACCATGCTCCACTCCATTTCTGTTCTGATTTGGACACCATGATATA 780  
Qy 798 GTTAAGTATCTGCTGCAAGAGTATTTGGAAGTTCAACCTCATGTTGTTAATATCTATGGA 857  
Db 781 GTTAAGTATCTGCTGCAAGAGTATTTGGAAGTTCAACCTCATGTTGTTAATATCTATGGA 840  
Qy 858 GATACCCCTTACACCTGGCATGCTCAATGGCAAAATTTGAAAGTTGCCAAGGAAATCATC 917  
Db 841 GATACCCCTTACACCTGGCATGCTCAATGGCAAAATTTGAAAGTTGCCAAGGAAATCATC 900  
Qy 918 CAATATACGAAACAGAAAGTCTGACTAAGGAAACATCTTCAGTGAACACGCTTTTCAT 977  
Db 901 CAATATACGAAACAGAAAGTCTGACTAAGGAAACATCTTCAGTGAACACGCTTTTCAT 960  
Qy 978 AGTCTTGTAACCTATGCAAGAGCATTTGACCTAGTCAAAATTTCTTCTGATCAGAAATGTC 1037  
Db 961 AGTCTTGTAACCTATGCAAGAGCATTTGACCTAGTCAAAATTTCTTCTGATCAGAAATGTC 1020  
Qy 1038 ATAAACATCAACCAAGAAAGGATGGGCAACCTGGAATTAACATCTGCTTGTCTACAC 1097  
Db 1021 ATAAACATCAACCAAGAAAGGATGGGCAACCTGGAATTAACATCTGCTTGTCTACAC 1080  
Qy 1098 GGTACATTCGCTGTTTCAAGTTCTTACTGGATAATGGAGCTGATATGAATCTAGTGGCT 1157  
Db 1081 GGTACATTCGCTGTTTCAAGTTCTTACTGGATAATGGAGCTGATATGAATCTAGTGGCT 1140  
Qy 1158 TGTGATCCGACAGGCTAGTGTGTAAGAGATGAGCAGACATGTTTGTATGTTGGGCTTAT 1217  
Db 1141 TGTGATCCGACAGGCTAGTGTGTAAGAGATGAGCAGACATGTTTGTATGTTGGGCTTAT 1200  
Qy 1218 GAAAGGGGATGATGCCATTTGCACTCTGAAAGCATTAAGAGACCAAGATGAA 1277  
Db 1201 GAAAGGGGATGATGCCATTTGCACTCTGAAAGCATTAAGAGACCAAGATGAA 1260  
Qy 1278 TTGCGCTGTAATGAATATTTCTACGCTGGAGAGATGGCTCTATGTTCTGTTCCATCA 1337  
Db 1261 TTGCGCTGTAATGAATATTTCTACGCTGGAGAGATGGCTCTATGTTCTGTTCCATCA 1320  
Qy 1338 CCCTTGGGGAAGATTAAGCATGACAAAAGAGAGGAGATATTTCTCTCTAGAGCT 1397  
Db 1321 CCCTTGGGGAAGATTAAGCATGACAAAAGAGAGGAGATATTTCTCTCTAGAGCT 1380  
Qy 1398 GGATGCTTACATTTCCATTTCTAGCTCTCAGABATTTAGTTCATCAGATTTATGSC 1457  
Db 1381 GGATGCTTACATTTCCATTTCTAGCTCTCAGAAATTTAGTTCATCAGATTTATGSC 1440

Qy 1458 TCAGGTTCTTTTGGGAAAGTATATAAAGCAGATGCAGAAATAAAATAGTGGCTATAAAA 1517  
Db 1441 TCAGGTTCTTTTGGGAAAGTATATAAAGCAGATGCAGAAATAAAATAGTGGCTATAAAA 1500  
Qy 1518 CGTTATCGAGCCAAATACCTACTGCTCCAAAGTCAGATGCGATATGTTTGGCCGAGGCTG 1577  
Db 1501 CGTTATCGAGCCAAATACCTACTGCTCCAAAGTCAGATGCGATATGTTTGGCCGAGGCTG 1560  
Qy 1578 TCCATTTCTGCGCAGCTCAATCATCCCTCGTAATTCAGTTTGTGGTGTCTTGTCTGAAT 1637  
Db 1561 TCCATTTCTGCGCAGCTCAATCATCCCTCGTAATTCAGTTTGTGGTGTCTTGTCTGAAT 1620  
Qy 1638 GATCCCAGCAGTTTCCCATTTGCTCAATCAATACATACAGGGGGTCTCTGTCTCCCTC 1697  
Db 1621 GATCCCAGCAGTTTCCCATTTGCTCAATCAATACATACAGGGGGTCTCTGTCTCCCTC 1680  
Qy 1698 CTTTCATGAGCAGAGAGGATTTCTTGATTTTGCAGTCTAAATTAATTTATTCAGTAGATGTT 1757  
Db 1681 CTTTCATGAGCAGAGAGGATTTCTTGATTTTGCAGTCTAAATTAATTTATTCAGTAGATGTT 1740  
Qy 1758 GCCAAAGGCATGAGATACCTTCAACCTGACACAGCCAAATTTATACATCGTGAATGAAC 1817  
Db 1741 GCCAAAGGCATGAGATACCTTCAACCTGACACAGCCAAATTTATACATCGTGAATGAAC 1800  
Qy 1818 AGTCACAAATATTTCTTCTCTATGAGGATGGCATGCTGTGGTGGCAGATTTTGGAGAAATCA 1877  
Db 1801 AGTCACAAATATTTCTTCTCTATGAGGATGGCATGCTGTGGTGGCAGATTTTGGAGAAATCA 1860  
Qy 1878 AGATTTCTACAGTCTCTGGATGAAGCAACATGACAAAAACAACTGGGAAACCTCGTTGG 1937  
Db 1861 AGATTTCTACAGTCTCTGGATGAAGCAACATGACAAAAACAACTGGGAAACCTCGTTGG 1920  
Qy 1938 ATGGCTCTGAGGTGTTCAACGAGTGCATCTCGGTACACCATCAAGCAGATGCTTTCAGC 1997  
Db 1921 ATGGCTCTGAGGTGTTCAACGAGTGCATCTCGGTACACCATCAAGCAGATGCTTTCAGC 1980  
Qy 1998 TATGCTCTGTGCTGTGGGAAATTTCTCACTGCGGAAATTTCCATTTCCCTCATCTCAAGCCA 2057  
Db 1981 TATGCTCTGTGCTGTGGGAAATTTCTCACTGCGGAAATTTCCATTTCCCTCATCTCAAGCCA 2040  
Qy 2058 GCGGCTCGGCGACGACATGCGTTACCAACCATCAGACCTCCCATTTGGCTTATTCATTT 2117  
Db 2041 GCGGCTCGGCGACGACATGCGTTACCAACCATCAGACCTCCCATTTGGCTTATTCATTT 2100  
Qy 2118 CCCAAGCCCATATCATCTCTGCTGATACGAGGTGGAGCGATGCTCTGCAACATTTGAGAGCC 2177  
Db 2101 CCCAAGCCCATATCATCTCTGCTGATACGAGGTGGAGCGATGCTCTGCAACATTTGAGAGCC 2160  
Qy 2178 GAATTTTCTGAACTTGTCTAAGTTAGAGAGTGTCTCTGCAACATTTGAGCTGATGCT 2237  
Db 2161 GAATTTTCTGAACTTGTCTAAGTTAGAGAGTGTCTCTGCAACATTTGAGCTGATGCT 2220  
Qy 2238 CTTGCTATCAAGTAACAGCAGTGGGTCTCTCACTTCTTCTTCTTCTGATTTGGCTGGTG 2297  
Db 2221 CTTGCTATCAAGTAACAGCAGTGGGTCTCTCACTTCTTCTTCTTCTGATTTGGCTGGTG 2280  
Qy 2298 AACCGGGAGGACCTCGCGGAGTCAATGTCGAGCAATTAAGAGTCTGTTTCCAAATGGAA 2357  
Db 2281 AACCGGGAGGACCTCGCGGAGTCAATGTCGAGCAATTAAGAGTCTGTTTCCAAATGGAA 2340  
Qy 2358 TATGCTCTAAATCAAGGTCCTATGCTGTTTGTGCCAAAGTGTGGACAAATTTCTCTCT 2417  
Db 2341 TATGCTCTAAATCAAGGTCCTATGCTGTTTGTGCCAAAGTGTGGACAAATTTCTCTCT 2400  
Qy 2418 CAAGGTCGTCTTTTGGAGAGATGAAAGAGTCTTCAATACACCCCATTTGACAAATAT 2477  
Db 2401 CAAGGTCGTCTTTTGGAGAGATGAAAGAGTCTTCAATACACCCCATTTGACAAATAT 2460  
Qy 2478 GGTATGATATCCCATGCCATGAGCTCAATGCAATTTTCAATTTTGGCGAAATAGTAGCAGC 2537  
Db 2461 GGTATGATATCCCATGCCATGAGCTCAATGCAATTTTCAATTTTGGCGAAATAGTAGCAGC 2520

QY	2538	TTTGGAGCAGCAGCTGACAGATTGCGCGTATACCTAAGGAGTGTCTTCCCGCAACT	2597	Db	181	AAATTACGCACTGAAATGGCTGTCTCTACTTCAATTAATGTTGCAATTTGTGGAGCAAG	240
Db	2521	TTTGGAGCAGCAGCTGACAGATTGCGCGTATACCTAAGGAGTGTCTTCCCGCAACT	2580	QY	288	AAATCAATATTCGAACCTTTATGTTGAAGGGCTCGGCCATCTCGACTGACAGAAT	347
QY	2598	GACAGCAACGATTCACCAACCGGCAAGCTGGCTTCCAACTATACATTTTACTCTCAAG	2657	Db	241	AAATCAATATTCGAACCTTTATGTTGAAGGGCTCGGCCATCTCGACTGACAGAAT	300
Db	2581	GACAGCAACGATTCACCAACCGGCAAGCTGGCTTCCAACTATACATTTTACTCTCAAG	2640	QY	348	GGATTTACAGCTTGGCAATTTAGCAGTTTACAGGATAATGCAAGATTGATCACTTCTCTG	407
QY	2658	GTCTCTTAATTTGGCTTGTCTTACTTGTCTTAAATCCCACTATTAGCAGCT	2717	Db	301	GGATTTACAGCTTGGCAATTTAGCAGTTTACAGGATAATGCAAGATTGATCACTTCTCTG	360
Db	2641	GTCTCTTAATTTGGCTTGTCTTACTTGTCTTAAATCCCACTATTAGCAGCT	2700	QY	408	CTTCAAGTGGAGCTGATATACAGCAGTTGATACGGTGGCTCACTGCCCTCCATATT	467
QY	2718	TTGGATTTGTGCTAAGCAATATATGCAAAAGAACCAAGACAAATGTATATGAAGAAT	2777	Db	361	CTTCAAGTGGAGCTGATATACAGCAGTTGATACGGTGGCTCACTGCCCTCCATATT	420
Db	2701	TTGGATTTGTGCTAAGCAATATATGCAAAAGAACCAAGACAAATGTATATGAAGAAT	2760	QY	468	GCTCAATAGCTGCCACCTAGAGGCTGTGATGTCTGTGCAACATGGAGCTAATGTC	527
QY	2778	TGTTTTTAATTTTGAATTAATAAATTTAGATCGTTACTTTGGAATGGAGCTTAAG	2837	Db	421	GCTCAATAGCTGCCACCTAGAGGCTGTGATGTCTGTGCAACATGGAGCTAATGTC	480
Db	2761	TGTTTTTAATTTTGAATTAATAAATTTAGATCGTTACTTTGGAATGGAGCTTAAG	2820	QY	528	AAATATTCAGATGCAGTTTTTTTCACTCCATTCATATTGCAGCGTACTATGGACATGAA	587
QY	2838	TCGTGGTGGACAGATAATATTTTCTCGGCTGAATATGTAGACTTGTGTTTG	2897	Db	481	AAATATTCAGATGCAGTTTTTTTCACTCCATTCATATTGCAGCGTACTATGGACATGAA	540
Db	2821	TCGTGGTGGACAGATAATATTTTCTCGGCTGAATATGTAGACTTGTGTTTG	2880	QY	588	CAGTAACTCGCCTTCTTTTGAATTTGGTGTGATGTAATGTAAGTGGTGAAGTTGGA	647
QY	2898	ACAGTATGGGTTTATTTCTTGAACATTTGTTCAATTTCTTTCTCAATATGTTCTCT	2957	Db	541	CAGTAACTCGCCTTCTTTTGAATTTGGTGTGATGTAATGTAAGTGGTGAAGTTGGA	600
Db	2881	ACAGTATGGGTTTATTTCTTGAACATTTGTTCAATTTCTTTCTCAATATGTTCTCT	2940	QY	648	GATAGACCTCCACCTAGCATCTGCAAAAGGATTTCTTGAATATTGCAAACTCTTTGATG	707
QY	2958	AGTGTTCACCTCTGTATTAAGATTCTTTGGTGAATAG	2997	Db	501	GATAGACCTCCACCTAGCATCTGCAAAAGGATTTCTTGAATATTGCAAACTCTTTGATG	660
Db	2941	AGTGTTCACCTCTGTATTAAGATTCTTTGGTGAATAG	2980	QY	708	GAAGAAGCAAGACAGATGTAATGCTCAAGATAATGAAGACATGTCCTCCACTCCAT	767
RESULT 6							
AX056437							
LOCUS							
DEFINITION							
Sequence 81 from Patent WO0073469.							
AX056437							
VERSION							
AX056437.1 GI:12229144							
KEYWORDS							
Homo sapiens (human)							
SOURCE							
Homo sapiens							
ORGANISM							
Homo sapiens							
REFERENCE							
AUTHORS							
TITLE							
JOURNAL							
FEATURES							
source							
1..2508							
/organism="Homo sapiens"							
/mol_type="unassigned DNA"							
/db_xref="taxon:9606"							
ORIGIN							
Query Match							
Best Local Similarity 82.9%; Score 2508; DB 6; Length 2508;							
Matches 2508; Conservative 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;							
QY	48	ATGGGAATTTAATCTAGACCAACCCAACTTGATGTAATGGAAGAAAAGTC	107	QY	1068	CACACTGGATTACACTCTGCTTCTACACGGTTCACATTCGCTGCTTTCAGTTCCTTACTG	1127
Db	1	ATGGGAATTTAATCTAGACCAACCCAACTTGATGTAATGGAAGAAAAGTC	60	Db	1021	CACACTGGATTACACTCTGCTTCTACACGGTTCACATTCGCTGCTTTCAGTTCCTTACTG	1080
QY	108	AGTGAATCATATGTTATCACAATAGAAGATTAGAATGACCTGCAGATCAAGGAAAA	167	QY	1128	GATAATGGAGCTGATATGTAATCTAGTGGCTTGTGATCCCAGCAGGTCTAGTGGTGA	1187
Db	61	AGTGAATCATATGTTATCACAATAGAAGATTAGAATGACCTGCAGATCAAGGAAAA	120	Db	1081	GATAATGGAGCTGATATGTAATCTAGTGGCTTGTGATCCCAGCAGGTCTAGTGGTGA	1140
QY	168	GAACCTGACAGAACTAAGGAATATTTGGCTCTGATGAAGCTTCAGTAAAGTCAATTTA	227	QY	1188	GATGAGCAGACATGTTTGTGATGGCTTATGAAAAGGGCATGATGCCATTTGTCACATC	1247
Db	121	GAACCTGACAGAACTAAGGAATATTTGGCTCTGATGAAGCTTCAGTAAAGTCAATTTA	180	Db	1141	GATGAGCAGACATGTTTGTGATGGCTTATGAAAAGGGCATGATGCCATTTGTCACATC	1200
QY	228	AATTACCCGACTGAAAATGGGCTGTCTCTACTTCATTTATGTTGCAATTTGTGGAGCAAG	287	QY	1248	CTGAAGCATTTAAGAGACCAAGATGAATTTGCCCTGTAATTTCTCAGGCTCGA	1307
Db				Db	1201	CTGAAGCATTTAAGAGACCAAGATGAATTTGCCCTGTAATTTCTCAGGCTCGA	1260
QY				QY	1308	GGAGATGGCTCTGATGTCTGTTCATCACCTTGGGGAAGATTAAAAGCATGACAAAA	1367

Db 1361 GGAGATGGCTCCATATGTTGCTGTTCCATCACCCCTTGGGGAAAGATTAAAGCATGACAAA 1320  
Qy 1368 GAGAAGCAGATATTTCTCTCTTAAGAGTGGATTGCTTCAATTTCCATTTCCAGTCT 1427  
Db 1321 GAGAAGCAGATATTTCTCTCTTAAGAGTGGATTGCTTCAATTTCCATTTCCAGTCT 1380  
Qy 1428 TCAGAAATTCAGTTCATGAGATATTTGGCTCAGGTTCTTTTGGAAAGATATATAAGGA 1487  
Db 1381 TCAGAAATTCAGTTCATGAGATATTTGGCTCAGGTTCTTTTGGAAAGATATATAAGGA 1440  
Qy 1488 CGATGCAAAATAAAATAGTGGCTATAAAACGTTATCGAGCCAAATACCTACTGCTCCCAAG 1547  
Db 1441 CGATGCAAAATAAAATAGTGGCTATAAAACGTTATCGAGCCAAATACCTACTGCTCCCAAG 1500  
Qy 1548 TCAGATGAGATATTTTCCGAGAGGTGTCATTTCTCTGCGAGCTCAATCATCCCTGCG 1607  
Db 1501 TCAGATGAGATATTTTCCGAGAGGTGTCATTTCTCTGCGAGCTCAATCATCCCTGCG 1560  
Qy 1608 GTAATTCAGTTTGGGTGCTTCTCTTGAATGATCCAGCCAGTTTGCCATTGCTCACTCA 1667  
Db 1561 GTAATTCAGTTTGGGTGCTTCTCTTGAATGATCCAGCCAGTTTGCCATTGCTCACTCA 1620  
Qy 1668 TACATATCAGGGGTCTCTGTTCTCTCCCTCTTCATGACAGAGGATCTTGAATTG 1727  
Db 1621 TACATATCAGGGGTCTCTGTTCTCTCCCTCTTCATGACAGAGGATCTTGAATTG 1680  
Qy 1728 CAGTCTAAATTAATTTAGTTCAGTAGATGTTGCCAAGGCATGGATACCTTACAACTG 1787  
Db 1681 CAGTCTAAATTAATTTAGTTCAGTAGATGTTGCCAAGGCATGGATACCTTACAACTG 1740  
Qy 1788 ACACAGCAATATATACATCTGACTTGAACAGTCACAAATATTTCTCTATGAGGATGG 1847  
Db 1741 ACACAGCAATATATACATCTGACTTGAACAGTCACAAATATTTCTCTATGAGGATGG 1800  
Qy 1848 CATGCTGTGTGGCAGATTTTGAGAAATCAAGATTTCTACAGTCTCTGATGAAGACAC 1907  
Db 1801 CATGCTGTGTGGCAGATTTTGAGAAATCAAGATTTCTACAGTCTCTGATGAAGACAC 1860  
Qy 1908 ATGCAAAAACACCTGGGAACCTCCGTTGGATGGCTCCCTGAGTGTTCACGAGTGCCT 1967  
Db 1861 ATGCAAAAACACCTGGGAACCTCCGTTGGATGGCTCCCTGAGTGTTCACGAGTGCCT 1920  
Qy 1968 CGGTACACCATCAAGCAGATGCTTACGATATGCTCTGTGTCTGTTGGAAATTTCTCACT 2027  
Db 1921 CGGTACACCATCAAGCAGATGCTTACGATATGCTCTGTGTCTGTTGGAAATTTCTCACT 1980  
Qy 2028 GGCAGAAATTCATTCGCTATCTCAAGCCAGCGCTCGCGCAGCAGATGCTTACCAC 2087  
Db 1981 GGCAGAAATTCATTCGCTATCTCAAGCCAGCGCTCGCGCAGCAGATGCTTACCAC 2040  
Qy 2088 CACATCAGACCTCCCATTTGGCTATTCCATTTCCCAAGCCCATATCATCTCTGCTGATACGA 2147  
Db 2041 CACATCAGACCTCCCATTTGGCTATTCCATTTCCCAAGCCCATATCATCTCTGCTGATACGA 2100  
Qy 2148 GGGTGGAAACGATGCTCTGAAGAGAGCCGAAATTTCTGAAGTTGTCTATGAAGTTAGNA 2207  
Db 2101 GGGTGGAAACGATGCTCTGAAGAGAGCCGAAATTTCTGAAGTTGTCTATGAAGTTAGNA 2160  
Qy 2208 GAGTGTCTCTGCAACATTCAGTGTATCTCTCTGCAATCAAGTAAACAGCAGTGGGTCTCTC 2267  
Db 2161 GAGTGTCTCTGCAACATTCAGTGTATCTCTCTGCAATCAAGTAAACAGCAGTGGGTCTCTC 2220  
Qy 2268 TCACCTTCTTCTTCTCTGATGCTGCTGTAACCGGGAGGACCTCGCGGAGTCAATGTG 2327  
Db 2221 TCACCTTCTTCTTCTCTGATGCTGCTGTAACCGGGAGGACCTCGCGGAGTCAATGTG 2280  
Qy 2328 GCACATTAAGAGTCTGTTTTCGAAATGCAATATGCTCTAAATTCGAGGTCCTATGCTGCT 2387  
Db 2281 GCACATTAAGAGTCTGTTTTCGAAATGCAATATGCTCTAAATTCGAGGTCCTATGCTGCT 2340  
Qy 2388 TTGTCCCAAGTCTGGACAAATATTCCTCTCAAGGTCGTGCTTTGGAGGAGATGAAAGA 2447  
Db 2341 TTGTCCCAAGTCTGGACAAATATTCCTCTCAAGGTCGTGCTTTGGAGGAGATGAAAGA 2400

Qy 2448 AGTCTTCAATACACACCCATTGACAAAATATGCTATGATCCGATCCCATGAGCTCAATG 2507  
Db 2401 AGTCTTCAATACACACCCATTGACAAAATATGCTATGATCCGATCCCATGAGCTCAATG 2460  
Qy 2508 CATTTTCATTTCTTCCGAAATAGTAGCAGCTTTGAGGACAGCAGCTGA 2555  
Db 2461 CATTTTCATTTCTTCCGAAATAGTAGCAGCTTTGAGGACAGCAGCTGA 2508

RESULT 7  
BD269560 2505 bp DNA linear PAT 17-JUL-2003  
LOCUS  
DEFINITION Novel CARX protein and nucleic acid molecules and uses therefor.  
ACCESSION BD269560  
VERSION BD269560.1 GI:33079328  
KEYWORDS JP 2002535962-A/2.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
1 (bases 1 to 2505)  
Raju,J.  
AUTHORS  
TITLE Novel CARX protein and nucleic acid molecules and uses therefor  
JOURNAL MILLENNIUM PHARMACEUTICALS INC  
COMMENT  
OS Homo sapiens (human)  
PN JP 2002535962-A/2  
PD 29-OCT-2002  
PF 10-DEC-1999 JP 2000586772  
PR 11-DEC-1998 US 60/111938, 14-APR-1999 US 09/291839 PI  
JEYASEELAN RAJU  
PC C12N15/09,C07K16/40,C12N1/15,C12N1/19,C12N1/21,C12N5/10,C12N9/  
10,C12Q1/02.  
PC C12Q1/68,G01N33/15,G01N33/50,G01N33/53,G01N33/68//A61K38/45,  
PC A61K39/395,  
PC A61K45/00,A61P9/04,A61P9/10,A61P35/00,A61P43/00,C12N15/00, PC  
C12N5/00,  
PC A61K37/52  
CC Novel CARX protein and nucleic acid molecules and uses CC  
therefor  
FH Key Location/Qualifiers  
FT CDS Location/Qualifiers  
1..2505  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"

ORIGIN  
Query Match 82.8%; Score 2505; DB 6; Length 2505;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 48 ATGGGAATTTATAAATCTAGACCCCAACCTGCTACTGATGAATGGAGAAAGATC 107  
Db 1 ATGGGAATTTATAAATCTAGACCCCAACCTGCTACTGATGAATGGAGAAAGATC 60  
Qy 108 AGTGAATCATATGTTATCAATAGAAAGATTAGAAGATGACCTGCAGATCAAGGAAAA 167  
Db 61 AGTGAATCATATGTTATCAATAGAAAGATTAGAAGATGACCTGCAGATCAAGGAAAA 120  
Qy 168 GAATGACAGAACTAGGAATATATTTGGCTCTGATGAAGCCTTCAGTAAAGTCAATTTA 227  
Db 121 GAATGACAGAACTAGGAATATATTTGGCTCTGATGAAGCCTTCAGTAAAGTCAATTTA 180  
Qy 228 AATTACCGCACTGAAATGGGCTGCTCTACTTCATTTATGTTGCAATTTGTGGAGCAAG 287  
Db 181 AATTACCGCACTGAAATGGGCTGCTCTACTTCATTTATGTTGCAATTTGTGGAGCAAG 240  
Qy 288 AAATTCACATATTCGAATCTTATGTTGAAAGGCTCCGCCATCTCGACTGACAGAAAT 347  
Db 241 AAATTCACATATTCGAATCTTATGTTGAAAGGCTCCGCCATCTCGACTGACAGAAAT 300



```
Db 2461 CATTTTCATTCTGCCGAATAGTAGCAGCTTTGAGGACGACG 2505

RESULT 8
LOCUS AR269112 2505 bp DNA linear PAT 10-APR-2003
DEFINITION Sequence 3 from patent US 6500654.
ACCESSION AR269112
VERSION AR269112.1 GI:29699955
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 2505)
AUTHORS Raju J.
TITLE CARP protein and nucleic acid molecules and uses therefor
JOURNAL Patent: US 6500654-A 3 31-DEC-2002;
FEATURES
    source
        /organism="unknown"
        /mol_type="genomic DNA"

ORIGIN
Query Match 82.8%; Score 2505; DB 6; Length 2505;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 48 ATGGGAATTAATACTAGACCAACCCAACTGTACTGATGATGGAAGAAAAGTC 107
Db 1 ATGGGAATTAATACTAGACCAACCCAACTGTACTGATGATGGAAGAAAAGTC 60

QY 108 AGTGAATCATATGTTATACAAATAGAAAGATTAGAAGTACCTGCAGATCAAGGAAAA 167
Db 61 AGTGAATCATATGTTATACAAATAGAAAGATTAGAAGTACCTGCAGATCAAGGAAAA 120

QY 168 GAATGACAGAACTAAGGATATATTTGGCTCTGATGAAGCCCTTCAGTAAAGTCAATTTA 227
Db 121 GAATGACAGAACTAAGGATATATTTGGCTCTGATGAAGCCCTTCAGTAAAGTCAATTTA 180

QY 228 AATTACCGCACTGAAATGGGCTGTCTACTACTTATTTGTCATTTGTGGAGGCAAG 287
Db 181 AATTACCGCACTGAAATGGGCTGTCTACTACTTATTTGTCATTTGTGGAGGCAAG 240

QY 288 AATTCATATTCGAATCTTTATGTTGAAGGGCTCGGCCCATCTCGACTGACAGAAAT 347
Db 241 AATTCATATTCGAATCTTTATGTTGAAGGGCTCGGCCCATCTCGACTGACAGAAAT 300

QY 348 GGATTTACAGCTTGCATTTAGCAGTTTACAAGGATAATGCAGAAATTGATCACTTCTCTG 407
Db 301 GGATTTACAGCTTGCATTTAGCAGTTTACAAGGATAATGCAGAAATTGATCACTTCTCTG 360

QY 408 CTTTCACAGTGGAGTGATATACAGCAGGTTGGATACGGTGGCTCACTCGCCCTCCATAT 467
Db 361 CTTTCACAGTGGAGTGATATACAGCAGGTTGGATACGGTGGCTCACTCGCCCTCCATAT 420

QY 468 GCTACATAGCTGCCACCTAGAGGCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 527
Db 421 GCTACATAGCTGCCACCTAGAGGCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480

QY 528 AATATTCAGAGTGCAGTGTGTTTTTCACTCCATATTCATATTCGACGCTACTATGGACATGAA 587
Db 481 AATATTCAGAGTGCAGTGTGTTTTTCACTCCATATTCATATTCGACGCTACTATGGACATGAA 540

QY 588 CAGTAACTGCCTTCTTTTGAATTTGGTGTGATGTAATGTAAGTGGTGAAGTTGA 647
Db 541 CAGTAACTGCCTTCTTTTGAATTTGGTGTGATGTAATGTAAGTGGTGAAGTTGA 600

QY 648 GATAGCCCTCCACCTAGCATCTGCAAAAGGATCTTGAATATTCGAAAATCTTTGATG 707
Db 601 GATAGCCCTCCACCTAGCATCTGCAAAAGGATCTTGAATATTCGAAAATCTTTGATG 660

QY 708 GAAGAGGCGCAAGCAGATGTAATGCTCAAGATAATGAAGACCATGTCCCACTCCAT 767
```

```
Db 661 GAAGAGGCGCAAGCAGATGTAATGCTCAAGATAATGAAGACCATGTCCCACTCCAT 720
QY 768 TTCTGTTCTCGATTTGGACACCATGATATAGTTAAGTATCTGCTGCAAAAGTGAATTGAA 827
Db 721 TTCTGTTCTCGATTTGGACACCATGATATAGTTAAGTATCTGCTGCAAAAGTGAATTGAA 780
QY 828 GTTCAACCTCATCTGTTAATATCTATGAGATACCCCTTACACCTGCGCATGCTACAAAT 887
Db 781 GTTCAACCTCATCTGTTAATATCTATGAGATACCCCTTACACCTGCGCATGCTACAAAT 840
QY 888 GGCAAAATTTGAAGTTCGCAAGGAAATCATCCAAATATCAGGAACAGAAAGTCTGACTAAG 947
Db 841 GGCAAAATTTGAAGTTCGCAAGGAAATCATCCAAATATCAGGAACAGAAAGTCTGACTAAG 900
QY 948 GAAAACATCTTCAGTGAACACAGCTTTTTCATAGTGTGTGCTATGCGCAAGAGCATGAC 1007
Db 901 GAAAACATCTTCAGTGAACACAGCTTTTTCATAGTGTGTGCTATGCGCAAGAGCATGAC 960
QY 1008 CTAGTCAAAATTTCTTCTTGATCAGAAATGTCTAAACATCAACCCCAAGGAAGGATGGG 1067
Db 961 CTAGTCAAAATTTCTTCTTGATCAGAAATGTCTAAACATCAACCCCAAGGAAGGATGGG 1020
QY 1068 CACTGGAATACACTCTGCTGTACACAGGTCAATTCGCTGGTTCAGTCTTACTG 1127
Db 1021 CACTGGAATACACTCTGCTGTACACAGGTCAATTCGCTGGTTCAGTCTTACTG 1080
QY 1128 GATAATGGAGCTCATATGAATCTAGTGGCTTGTGATCCCGACAGGTCTAGTGGTGA AAAA 1187
Db 1081 GATAATGGAGCTCATATGAATCTAGTGGCTTGTGATCCCGACAGGTCTAGTGGTGA AAAA 1140
QY 1188 GATGAGCAGACATGTTTGTATGTTGGCTTATGAAAAGGCGATGATGCCATTTGTCACTC 1247
Db 1141 GATGAGCAGACATGTTTGTATGTTGGCTTATGAAAAGGCGATGATGCCATTTGTCACTC 1200
QY 1248 CTGAAGCATTTAAGAGACCAAGATGAATGTCCTGTAATGAATATTTCTCAGCTGGA 1307
Db 1201 CTGAAGCATTTAAGAGACCAAGATGAATGTCCTGTAATGAATATTTCTCAGCTGGA 1260
QY 1308 GGAGATGGCTCTCTATGTTGTTCCATCACCTTTGGGGAAGATTAAAAGCATGACAAA 1367
Db 1261 GGAGATGGCTCTCTATGTTGTTCCATCACCTTTGGGGAAGATTAAAAGCATGACAAA 1320
QY 1368 GAGAGGCGAGATATTTCTCTCTAAGAGCTGATGCTTCCATTTCCATTTCCATTTCCATTT 1427
Db 1321 GAGAGGCGAGATATTTCTCTCTAAGAGCTGATGCTTCCATTTCCATTTCCATTTCCATTT 1380
QY 1428 TCAGAAATTCAGTTCCATGAGATTATTTGGCTCAGGTTCTTTTGGGAAAGTATATAAGGA 1487
Db 1381 TCAGAAATTCAGTTCCATGAGATTATTTGGCTCAGGTTCTTTTGGGAAAGTATATAAGGA 1440
QY 1488 CGATGCAAGAAATAAATAGTGGCTATAAAAAGTTATCGAGCAATACCTTACTGCTCCAG 1547
Db 1441 CGATGCAAGAAATAAATAGTGGCTATAAAAAGTTATCGAGCAATACCTTACTGCTCCAG 1500
QY 1548 TCAGATGTGATATGTTTTCGAGAGGTTGTCATTTCTGCGAGCTCAATCATCCCTGTC 1607
Db 1501 TCAGATGTGATATGTTTTCGAGAGGTTGTCATTTCTGCGAGCTCAATCATCCCTGTC 1560
QY 1608 GTAAATTCAGTTTGTGGTGTCTGTTGAATGATGCCAGCAGTTTGGCCATTTGTCTCACTAA 1667
Db 1561 GTAAATTCAGTTTGTGGTGTCTGTTGAATGATGCCAGCAGTTTGGCCATTTGTCTCACTAA 1620
QY 1668 TACATATCAGGGGTTCTGTTCTCCCTTCTCATGAGCAAGAGGATTTCTTGAATTG 1727
Db 1621 TACATATCAGGGGTTCTGTTCTCCCTTCTCATGAGCAAGAGGATTTCTTGAATTG 1680
QY 1728 CAGTCTAAAATTAATTTAGCAGTAGATGTTGCAAGGATGAGTACCTTCAACACCTG 1787
Db 1681 CAGTCTAAAATTAATTTAGCAGTAGATGTTGCAAGGATGAGTACCTTCAACACCTG 1740
QY 1788 ACACAGGCAATTTATCATGCTGAGCTTGAACAGTCAATATTTCTTCTTATGAGGATGG 1847
```

[illegible]

```
QY 1128 GATAATGGAGCTGATATGAATCTAGTGGCTTGTGATCCCGAGGAGTCTAGTGGTGAATAA 1187
Db 1081 GATATGGAGCTGATATGAATCTAGTGGCTTGTGATCCCGAGGAGTCTAGTGGTGAATAA 1140
QY 1188 GATGAGCAGACATGTTTGTATGGGCTTATGAAAAAGGAGCATGATGCCATTGTCAACATC 1247
Db 1141 GATGAGCAGACATGTTTGTATGGGCTTATGAAAAAGGAGCATGATGCCATTGTCAACATC 1200
QY 1248 CTGAAGCATTTAAGAGACACAGATGAATTCGCCCTGTAAATGATATTTCTCAGCCTGGA 1307
Db 1201 CTGAAGCATTTAAGAGACACAGATGAATTCGCCCTGTAAATGATATTTCTCAGCCTGGA 1260
QY 1308 GGAGATGGCTCCATGATGTCTGTTCATACCCCTGGGGAAGATTAAAAAGCATGACAAA 1367
Db 1261 GGAGATGGCTCCATGATGTCTGTTCATACCCCTGGGGAAGATTAAAAAGCATGACAAA 1320
QY 1368 GAGAAGCAGATATTTCTCTCCTAAGAGCTGGATTCCTTCATATTCCTCAGCTC 1427
Db 1321 GAGAAGCAGATATTTCTCTCCTAAGAGCTGGATTCCTTCATATTCCTCAGCTC 1380
QY 1428 TCAGAAATTCAGTTCCATGATATTTGGCTCAGGTTCTTTGGGAAGTATATAAGGA 1487
Db 1381 TCAGAAATTCAGTTCCATGATATTTGGCTCAGGTTCTTTGGGAAGTATATAAGGA 1440
QY 1488 CGATCGAATAATAAATAGTGGCTATAAAACGTTATCGAGCCAATACCTACTGCTCCAAG 1547
Db 1441 CGATCGAATAATAAATAGTGGCTATAAAACGTTATCGAGCCAATACCTACTGCTCCAAG 1500
QY 1548 TCAGATGTGATATGTTTCCGAGAGGTGCTCAATCTCTGCCAGCTCAATCATCCCTGC 1607
Db 1501 TCAGATGTGATATGTTTCCGAGAGGTGCTCAATCTCTGCCAGCTCAATCATCCCTGC 1560
QY 1608 GTAAATTCAGTTTGGGTGCTGCTGTAATGATCCAGCAGTTTCCCATTTGCTCACTCAA 1667
Db 1561 GTAAATTCAGTTTGGGTGCTGCTGTAATGATCCAGCAGTTTCCCATTTGCTCACTCAA 1620
QY 1668 TACATATCAGGGGTTCTCTGTTCTCCCTCTCATGAGCAGAGGATTTCTGATTG 1727
Db 1621 TACATATCAGGGGTTCTCTGTTCTCCCTCTCATGAGCAGAGGATTTCTGATTG 1680
QY 1728 CAGTCTAAATTAATTAATGTCAGTAGATGTTGCCAAAGGATGAGTACCTTCACAACTG 1787
Db 1681 CAGTCTAAATTAATTAATGTCAGTAGATGTTGCCAAAGGATGAGTACCTTCACAACTG 1740
QY 1788 ACACAGCAATTAATACATCGTACTGAAACAGTCAACAATATTTCTCTATCAGGATGG 1847
Db 1741 ACACAGCAATTAATACATCGTACTGAAACAGTCAACAATATTTCTCTATCAGGATGG 1800
QY 1848 CATGCTGTGTGGCAGATTTGGAGATCAAGATTTCTACAGTCTGTGATGAAGACAC 1907
Db 1801 CATGCTGTGTGGCAGATTTGGAGATCAAGATTTCTACAGTCTGTGATGAAGACAC 1860
QY 1908 ATGCAAAAACAACTGGGAACCTCCGTTGGATGGCTCCTGAGTGTTTCAGCAGTGCACT 1967
Db 1861 ATGCAAAAACAACTGGGAACCTCCGTTGGATGGCTCCTGAGTGTTTCAGCAGTGCACT 1920
QY 1968 CGGTACACCATCAAGCAGATGCTTCAGCTATGCTCTGTGTCTGTGGAAATTTCTACT 2027
Db 1921 CGGTACACCATCAAGCAGATGCTTCAGCTATGCTCTGTGTCTGTGGAAATTTCTACT 1980
QY 2028 GGCGAAATTCATTCGCTCATCTCAAGCAGCGGCTGGCGAGCAGACATGGCTTACCAC 2087
Db 1981 GGCGAAATTCATTCGCTCATCTCAAGCAGCGGCTGGCGAGCAGACATGGCTTACCAC 2040
QY 2088 CACATCAGACTCCCATTTGGCTATTCATTCCTCAAGCCCATATCACTCTCTGTGATACGA 2147
Db 2041 CACATCAGACTCCCATTTGGCTATTCATTCCTCAAGCCCATATCACTCTCTGTGATACGA 2100
QY 2148 GGGTGGAAACCATGCTCTGAAGAGACCGCAATTTCTGAAGTTGCTCATGAAGTTAGAA 2207
Db 2101 GGGTGGAAACCATGCTCTGAAGAGACCGCAATTTCTGAAGTTGCTCATGAAGTTAGAA 2160
```

```
QY 2208 GAGTGTCTCTGCAACATTTGAGCTGATGTCTCTGTCATCAAGTAAACAGAGTGGTCTCTC 2267
Db 2161 GAGTGTCTCTGCAACATTTGAGCTGATGTCTCTGTCATCAAGTAAACAGAGTGGTCTCTC 2220
QY 2268 TCACCTTCTTCTTCTTCTGATTTGCTGCTGTAACCCGGGAGGACCTGGCCGGAGTCAATG 2327
Db 2221 TCACCTTCTTCTTCTTCTGATTTGCTGCTGTAACCCGGGAGGACCTGGCCGGAGTCAATG 2280
QY 2328 GCAGCATTAAGAAGTCTGTTTCGAATTTGGAATATGCTCTAAATGCAAGGTCCTATCTGCT 2387
Db 2281 GCAGCATTAAGAAGTCTGTTTCGAATTTGGAATATGCTCTAAATGCAAGGTCCTATCTGCT 2340
QY 2388 TTGTCCCAAGTCTCGCAATATTTCTCTCAAGTCTCTTTTGGAGGAGATGAAAAGA 2447
Db 2341 TTGTCCCAAGTCTCGCAATATTTCTCTCAAGTCTCTTTTGGAGGAGATGAAAAGA 2400
QY 2448 AGTCTTCAATACACACCCATTTGACAAATATGCTATGCTATGATCCGATCCCATGAGCTCAATG 2507
Db 2401 AGTCTTCAATACACACCCATTTGACAAATATGCTATGCTATGATCCGATCCCATGAGCTCAATG 2460
QY 2508 CATTTTCATTTCTGCGAAATAGTAGCAGCTTTGAGGACAGCAGC 2552
Db 2461 CATTTTCATTTCTGCGAAATAGTAGCAGCTTTGAGGACAGCAGC 2505

RESULT 10
BD269561 3026 bp DNA linear PAT 17-JUL-2003
LOCUS Novel CARP protein and nucleic acid molecules and uses therefor.
DEFINITION
ACCESSION BD269561
VERSION BD269561.1 GI:33079329
KEYWORDS JP 2002535962-A/3.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 3026)
AUTHORS Raju,J.
TITLE Novel CARP protein and nucleic acid molecules and uses therefor
JOURNAL Patent: JP 2002535962-A 3 29-OCT-2002;
MILLENNIUM PHARMACEUTICALS INC
COMMENT OS Rattus norvegicus (rat)
PN JP 2002535962-A/3
PD 29-OCT-2002
PF 10-DEC-1999 JP 2000586772
PR 11-DEC-1998 US 60/111938,14-APR-1999 US 09/291839 PI
JEVASEELAN RAJU
PC C12N15/09,C07K16/40,C12N1/15,C12N1/19,C12N1/21,C12N5/10,C12N9/
PC 10,C12Q1/02.
PC C12Q1/68,G01N33/15,G01N33/50,G01N33/53,G01N33/68//A61K38/45,
PC A61K39/395,
PC A61K45/00,A61P9/04,A61P9/10,A61P35/00,A61P43/00,C12N15/00, PC
C12N5/00,
PC A61K37/52
CC Novel CARP protein and nucleic acid molecules and uses CC
therefor
FH Key Location/Qualifiers
FT CDS Location/Qualifiers
1..3026
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"

ORIGIN
Query Match 66.4%; Score 2010; DB 6; Length 3026;
Best Local Similarity 80.6%; Pred. No. 0;
Matches 2425; Conservative 0; Mismatches 570; Indels 13; Gaps 6;

QY 16 CGGCCCTGGAGAAAGGAAGAACTTATAATAATGGGAATTTATAAATCTAGCAACCC 75
Db 29 CAGCACAGGAGAAAGCAAAAGACTCTCTTTAAATGGGAATTTACAAATCCAGCAACAC 88
```

QY 76 AAACCTTGACTGATGAATGGAGAAAAAGTCAAGTCAATATGTTATCACAATAGAAA 135  
DB 89 AGACTTGTCTTGAATGGAAGAGAAAGTTAGTGAATCTTACGCTATTATCATAGAAA 148  
QY 136 GATTAGAGATCAACCTGCAGATCAAGGAAAAAGAACTGACAGAACTAAGGAAATATTTG 195  
DB 149 GCTGGAGGATCAACCTGCAGATCAAGGAAATGAATTTCAAGAACTAAGGCAATCTTG 208  
QY 196 GCTCTGATGAACCTTCACTAGTAAAGTCAATTTAAATACCGCACTGAAAAATGGGCTGTCTC 255  
DB 209 GCTCTGATGAAGCCTTCACTGAAAGTCAAGTCAATTTAAATACCGCAAGCTGGGCTGTCCC 268  
QY 256 TACTTCATTTATGTTGCAATTTGCGAGGCAAGAAATCACAATATTCGAATCTTTATGTTGA 315  
DB 269 TGCTACACTCTGCTGTGTCTGTGGCGGCAACAAGTCAATATCCGTGCGCTTATGTTAA 328  
QY 316 AAGGGCTCGGCCCTCTCGACTGACAGAAAATGGATTTACAGCCTTGCAATTTAGCAGTTT 375  
DB 329 AAGGGCTCGGTCCATCCAGACTGACGAGAAATGGTTTCCAGCTCTGCACCTGGCGGTTT 388  
QY 376 ACAAGGATAATGCAGAAATGATCACTTCTGCTTCAAGTGGAGCTGATATACACGAG 435  
DB 389 ACAAGGACAGCCGGAATTTATCACTTCACTGTTGCACGCGAGCAGATGTTTCAGCAAG 448  
QY 436 TTGGATAGGGTGGCCTCACTGCCCTCCATATTTGCTACAATAGCTGGCCACTAGAGGCTG 495  
DB 449 TGGGATAGGGTGGCCTCACAGCCTCCACATAGCTGCAATAGCTGGACACCCAGAGGCTG 508  
QY 496 CTGATGTGCTGTGCAACATGAGAGCTAATGTCAATATTTCAAGATGCAAGTTTTTTCACTC 555  
DB 509 CAGAAATGTGCTACAACATGGGGCCAAATGTGAATGTTCAAGATGCGGCTTCTTCAACC 568  
QY 556 CATTGCATATTCACGCTACTATGACATGAACAGGTAACCTCGCCTTCTTTTGAATTTG 615  
DB 569 CACTGCATNTGCAACCTACTATGGGCACGAGAGGTAACAGGTGCTTTTGAAGTTG 628  
QY 616 GTGCTGATGTAATGTAAGTGGTGAAGTTGGAGATAGACCCCTCCACTAGCATCTGCAA 675  
DB 629 GTGCTGATGTAATGTAAGCGGTGAAGTTGGGACAGGCTCTGCACCTGGCCTCTGCAA 688  
QY 676 AAGGATCTTGAATATTCAAACTCTTGATGGAGAGGACAGCAAGCAGATGTGAATG 735  
DB 689 AGGGCTTCTTCAACATTTGTGAACCTCTGTGTAGAGAGGAGGAGCAAGCAGATGTGAACG 748  
QY 736 CTCAGAGATAATGAAGACCATGTCCCACTTCTGTTCTCGATTTGGACACCATGATA 795  
DB 749 CTCAGGACAAATGAAGACCACTGCCCTCTGCACCTTCTGTTCTCGATTTGGACACCAATA 808  
QY 796 TAGTTAAGTATCTGCTGCAAGTGAATTTGGAAGTTCAACCTCATGTTGTTAATATCTATG 855  
DB 809 TAGTGAGTACTGCTGCTCAGAGTGAATTTAGAGGTCACGCTCAGCTCAATTAACATCTATG 868  
QY 856 GAGATACCCCTTACACCTGGCATGTACAAATGGCAAAATTTGAAGTTGCCAAGGAAATCA 915  
DB 869 GTGACACTCTTGCACCTGGCATGCTACATGGAATTTTGAAGTTGCCAAGGAAATG 928  
QY 916 TCCAAATATCAGGAACAGAAAGTCTGACTAAGGAAACATCTTCAGTGAAACAGCTTTTC 975  
DB 929 TCCAGGTAACAGGAACGAAAGTCTGACTAAGGAAACATCTTCAGCGAGACAGCTTTTC 988  
QY 976 ATAGTGTGTTACCTATGGCAAGAGCATTGACCTTAGTCAAAATTTCTTTGATCAGAAATG 1035  
DB 989 ACAGTGTGTTACCTATGGCAAGAACATTTGACCTGGTCAAAATTTCTTTGATCAGAAATG 1048  
QY 1036 TCATAAATCAATCAACCAAGGAGGAGTGGGCACTGGAATTAACCTCTGCTGCTGTAC 1095  
DB 1049 CTGTGAACATTAACCAAGGAGAGATGGGCAACAGGATTTGCACTCTGCTGTGTAC 1108  
QY 1096 ACGGTTCATTCGCTGTTTCAAGTTCTTACTGGATAATGGAGCTGATATCAATCTAGTGG 1155  
DB 1109 ACGGCCATATCGGCCCTTGGTTTCAAGTTCTTACTTGTATTTGGTCAGATATGAATCTTGTG 1168

QY 1156 CTTGTGATCCAGCAGGTCTAGTGGTGAAGAAAGATGAGCAGACATGTTTGTATGGGCTT 1215  
DB 1169 CTTGTGATCCAGCAGGTCTAGTGGTGAAGAAAGATGAGCAGACATGTTTGTATGGGCTT 1228  
QY 1216 ATGAAAAAGGCGCATGATGCCATTTGTCACTCTCTGAAGCATTTATAGAGACCACAGATG 1275  
DB 1229 ACAGAAAGGACATGATGCCATTTGTACACTCTCTGAAGCACTACAAGACACCCAGGAGG 1288  
QY 1276 AATTGGCCCTCTAATGATATTTCTCAGCTGGAGGATGCTCCTATGTTGTTCTCCAT 1335  
DB 1289 AGCTGCCATGTAAAGAAATTTCCAGCTGGAGGAGATGCTCCTATGTTGTTCTTCTT 1348  
QY 1336 CACCTTTGGGAAAGATTTAAAGCATGACAAAAGAGAGGAGATATTTCTCTCTCTAAAG 1395  
DB 1349 CCCCCTTTGGCAAGATTTAAAGCATGACAAAAGAGAGGAGATGTTCTCTCTCTGAGG 1408  
QY 1396 CTGATTTGCCCTTCACTTTCCATCTTCACTCTCAGCTCTCAGAAATTTGATTTCCATGATATG 1455  
DB 1409 CTGAATACCTCTCCGCTTCCATCTCCTCACTCTCCGAAATTCAGTTTCCACGAGATATG 1468  
QY 1456 GCTCAGGTTCTTTTGGAAAGTATATAAGGAGCAGTGCAGAAATATAAATAGTGGCTATAA 1515  
DB 1469 GCTCGGTTCTTTTGGAAAGTCTATAAAGGCGATGCAAGAAATAAATAGTGGCGATCA 1528  
QY 1516 AACGTTATCGAGCAATACCTACTGCTCCAAAGTCAGATGTGATATGTTTTCGAGAGG 1575  
DB 1529 AACGATACCGAGCAACACCTACTGCTCCAAGTCAGACGTGGATATGTTTTCGAGAGG 1588  
QY 1576 TGTCCATCTCTGCGAGCTCAATCATCTCCCTGCTTAATTCAGTTTGTGGGTGCTGCTTGA 1635  
DB 1589 TGTCCATCTCTGCGAGCTCAACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1648  
QY 1636 ATGATCCAGCAGGTTTGCATTTGCTCAATACATATACATATCAGGGGTTCTCTGTTCTCC 1695  
DB 1649 ATGACCCAGTCAAGTTTGCATTTGCTCACTCAGTACATTTTCAGAGGCTCCCTGTTCTCC 1708  
QY 1696 TCCTTCATGACGAGAGAGGATTTCTTGATTTGAGCTTAAATTAATTTGTCAGTAGATG 1755  
DB 1709 TGCTTCATGAACAGAGAGAAATTTCTTGACTTGCACTTGAATTTGACCCAGCAATCATAC 1768  
QY 1756 TTGCAAAAGGCATGGAGTACCTTTCAAACTGCACACAGCCAAATTTATACATCGTACTTGA 1815  
DB 1769 TTGCAAGGCGATGGAGTACCTGCACAGCTTTGACCCAGCAATCATACCCGAGCTGA 1828  
QY 1816 ACAGTCAATATTTCTCTATGAGGATGGGATGCTGCTGCTGAGATTTTGGAGAT 1875  
DB 1829 ACAGCCAAATTTCTGCTCTATGAGGATGGCCATGCTGCTGCTGAGATTTTGGAGAT 1888  
QY 1876 CAAGATTTCTACAGCTCTCTGATGAAGACAAACATGACAAAACAACTGGGAACTCCGCT 1935  
DB 1889 CAAGATTTCTGAGTCCCTGATGAAGACAAACATGACAAAGCAGCAGGAACTGCGCT 1948  
QY 1936 GGATGGCTCTGAGGTTTCAACGAGTGCACTCGGTACACCATCAAGCAGATGTTCTTCA 1995  
DB 1949 GGATGGCCCTTGAGTGTTCACACAGTGCACGAGATACACCATCAAGGCTGATGTTCTCA 2008  
QY 1996 GCTATGCTGTGCTGTGGGAAATTTCTCACTGGGAAATTTCACTTCGCTCATCTCAAGC 2055  
DB 2009 GTTACTCCCTGTGCTGTGGGAGCTCTCTCTGAGAAATTTCCATTCGCTCATCTCAAGC 2068  
QY 2056 CAGCGTGGCGCAGCAGACATGGCTTACCAACATCAGAGCTCCCATTTGCTATTTCA 2115  
DB 2069 CAGCGCTGCGCAGCAGACATATGGCTATCACACATCAGACCGGCTATTTCA 2128  
QY 2116 TTCCAGAGCCATATCATCTCTGATAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2175  
DB 2129 TCCCAAGCCATCTCATCTCTGATACGGGCTTGGAAATGCAATGCTCTGAGGAGCAGC 2188  
QY 2176 CCGAATTTCTGAAGTTGTCAATGAAGTTAAGAGAGTGTCTCTGCAACATTTGAGCTGATG 2235  
DB 2189 CAGAGTTCTCTGAAGTCTGTAGCAACTGGAGAGTGGCTATGCAATTTGAGAGCTCATGT 2248  
QY 2236 CTCTGTCATCAAGTAACAGCAGTGGGCTCTCTCACCTTCTTCTTCTTCTGATTCCTGG 2295

Db	2249	CTCAGCATCAAGTAACAGCAGTGGGTCTCTGTGCACCTCTCTCTTCTCCGATTCGCTGC	2308
Qy	2296	TGAACCGGGAGACCTGGCCGGAGTCATGTGCAGCATTAGAAGTCGTTTCGAATTCG	2355
Db	2309	TGAGCCGGGAGGGCTTGGCCGAGGCCACGTGCAGCCTTACGAGCCGTTTGTAGTTGG	2368
Qy	2356	AATATGCTTAAATCAAGGTCTATGTGCTTTGTCCCAAGTGTGCAACAATATTCCT	2415
Db	2369	AGTATGCCCTAAATCAAGGTCTATGTGCTGGTGGTCCCAAGTGTGGAACACACTCA	2428
Qy	2416	CTCAAGGTCTGCTTTGGAGGAGATGAAGAAGTCTTCAATACACACCAATTCACAAT	2475
Db	2429	ATCCGGCCCTGTCTTTGGAGGAGATGAATAGGAGCACCAGTATTCACCTGTTGACAAAT	2488
Qy	2476	ATGGCTATGTATCCGATCCCATGAGCTCAATGCAATTTTCAATCTTCGCCAAATAGTAGCA	2535
Db	2489	ACGGCTATGTCTGTATCCCATGAGCTGACGACCTTCACTCCGCCAAGACGACAGCA	2548
Qy	2536	GCCTTGAGGACAGCAGCTGACAGCATTCGGCGTATACCTTAAGAGAGATTTTTTCCCGAA	2595
Db	2549	ACTTTGAGGACACCACTGACAG-GTCTGGCATACACCTTAAGGGGGGTCTCCCATCAGG	2607
Qy	2596	CTGACAGCAACGATTCACACACGCGAAGCTGGCTTCCAACTATAACATTTTACTCTCAA	2655
Db	2608	CTGACAGCAGTGAATTTACCCATGGCAGGCTTCTTCCAAATATAACGCCCTGCCCTCTG	2667
Qy	2656	AGGCTCTCTTAAATGGGCTTGTATTTTACCTGTCTTATTTAATTCGCCACTATGACAGG	2715
Db	2668	AGGT-TTCTTCAAATCGTCTGTCTTATTTCTAAGCTCGTTTAAATTCCTTCTACAGACAG	2726
Qy	2716	CTTTGGATTTGCTCCTTAAGGAATAATGCAAAAGAACCAAGACAGAATGTATCAAGA	2775
Db	2727	GCCTTGACTCATGCC--AAGCCTGAAGTGTCAAGAGCAGATACAGAATGTGCATGAGGA	2784
Qy	2776	ATTGTTTTTAATTTGTAAATTAATAAATAAATTTAGATCGTTACTTGGAAATGGAGCCTA	2835
Db	2785	ATTGTTCTTAGTTTGATATTAAAGCCCTTAATTGCTGGGCTGGGGTTCAAA----TC	2840
Qy	2836	AGTCTGTGGTGGACAGATAATAATTATGTTTTCTCGGCTGGAATATGTAGACTTCGTGT	2895
Db	2841	TGTGTAGATAGCTGGTGTACCCTTATGTATTTGTGTAGCAACAATCTGTGGGCTTGTGT	2900
Qy	2896	TGACAG-CTATGGGTTATTTCTTAGAACAATGTTTCATTTTCTTCATTATGTTACT	2954
Db	2901	TGAGGGTCTCCTGTGGGTTTCTTAAACCAAGCTGGCTGATTTATCTCCTGTGCTTT	2960
Qy	2955	TCTAGTGTTCACCTCTGTGATTAAGATCTTTGGTGAATAGAAAAAATAAAAAA	3014
Db	2961	-----GTGTGTACTTCTGTGATTAAGTCTCTTCGGTGATCTAGAAAAAATAAAAAA	3016
Qy	3015	AGGCGCGC 3022	
Db	3017	GGCGGCC 3024	
RESULT 11			
AR269113			
LOCUS			
SEQUENCE 7 from patent US 6500654.			
AR269113			
AR269113.1			
GI:29699956			
Unknown.			
Unknown.			
Unclassified.			
1. (bases 1 to 3026)			
Raju, J.			
CARK protein and nucleic acid molecules and uses therefor			
Patent: US 6500654-A 7 31-DEC-2002;			
Location/Qualifiers			
1. .3026			
/organism="unknown"			
source			



SOURCE	Unknown.
ORGANISM	Unknown.
REFERENCE	Unclassified.
AUTHORS	Raju, J.
TITLE	CARK protein and nucleic acid molecules and uses therefor
JOURNAL	Patent: US 6660490-A 7 09-DEC-2003;
FEATURES	Location/Qualifiers
source	1..3026
	/organism="unknown"
	/mol_type="genomic DNA"
ORIGIN	
	Query Match 66.4%; Score 2010; DB 6; Length 3026;
	Best Local Similarity 80.6%; Pred.No. 0;
	Matches 2425; Conservative 0; Mismatches 570; Indels 13; Gaps 6;
QY	16 CGGCCTCGAGAAAGGAGAAACTTATATAAATGGGAATTATAAATCTAGCAACCC 75
Db	29 CAGCACAGGAGAAAGCAAAGACTTCCTTTAAATGGGAATTACAATCCAGACCACAC 88
QY	76 AAACCTTGACTCATGAATGGAGAAGAAAAGCTCAGTGAATCATATGTGTATCACAATAGAA 135
Db	89 AGACCTTGCTTGATGAATGGAGAAGAAAGTGTAGTGAATCTTACGCTATTATCATAGAAA 148
QY	136 GATTAGAAGTAGCCTGCAGATCAAGAAAAAGAACTGACAGAACTAAGGAATATATTG 195
Db	149 GGCTGGAGGATAAACCTGCAGATCAAGAAAAATGAATTCAGAATTAAGGCACATCTTG 208
QY	196 GCTCTGATGAGGCCCTTCAGTAAAGTCAATTTAAATTTACCGCACTGAATGGCTGTCTC 255
Db	209 GCTCTGATGAGGCCCTTCAGTGAAGTCAGTTTTAAATTTACCGCACAGAGCGTGCCTGTCCC 268
QY	256 TACTTCATTTATGTTGCAATTTGGAGGCAAGAAATCACATATTCGAATCTCTTATGTTGA 315
Db	269 TGCTACACCTCTGCTGTGCTGTGGGGCAACAAGTCACATATCCGTGCGCTTAATGTTAA 328
QY	316 AAGSGCTCGGCCATCTCGACTGACAGAAATCGATTTACAGCCTTGCATTTAGCAGTTT 375
Db	329 AAGSGCTCGGTCATTCAGACTGACAGAAATGGGTTTTCCAGCTTCGACCTGGCCGGTTT 388
QY	376 ACAAGGATAATCGAGAATTGATCATCTTCTCTGCTTCACAGTGGAGTGANATACAGCAGG 435
Db	389 ACAAGGACAGCCCGAACTTATCACTTCACCTGTTGCACAGCGGACAGATGTTACGAAAG 448
QY	436 TTGGATACGGTGGCCCTCACTGCCCTCCATATTGCTACAATAGCTGGCCACCTAGAGCGTG 495
Db	449 TGGGATACGGTGGCCCTCACAGCCCTCCACATAGCTGCAATAGCTGGACACCCACAGAGCTG 508
QY	496 CTGATGTGCTGTTGCAACATGGAGCTAAATGCTCAATATTCAGATGCAGTTTTTTTCACTC 555
Db	509 CAGAACTGCTGCTACACATGGGGCCAATGTGAATGTTCAAGATCGCTTCTTTCACCC 568
QY	556 CATTCGATATTGGACGGTCACTATGGACATGAACAGGTAACTCGCTTCTTTTGAANAATTG 615
Db	569 CACTCGCAITTTCAAACATTTGAAACTCTATGGGCACAGCAGGTAAACAGTGTCTTTTGAAGTTTG 628
QY	616 GTGCTGATGTAATGTAAAGTGGTGAAGTTGGAGATAGACCCCTCCACCTAGCATCTGCAA 675
Db	629 GTGCTGATGTCAATGTAAGCGGTGAAGTTGGGACAGGCCCTCTGCACCTGGCCCTTGCAA 688
QY	676 AAGGATTTCTGAATATTGCAAACTCTTGATGGAAGAGCGAGCAAAAGCATGTGAATG 735
Db	689 AGGGCTTCTTCAAACATTTGAAACTCTGTGGTGAAGAGGGAGCAAGCAGATGTGAACG 748
QY	736 CTCGAAGATAATGAAGACCATGTGCCACTCCCATTTCTGTTTCTCGATTTGGACACCATGATA 795
Db	749 CTCGAGCAATGAAGACCAACGTCCCTCTGCACTTCTGTCTCGATTTGGACACCAATA 808
QY	796 TAGTTAAGTATCTCGCTGCAAGATGATTTGGAAGTTCAACTCATGTTGTTAATCTATG 855
Db	809 TAGTCAGCTACTCTCTCCAGATGACTTTAGAGTTCAGGCTTCAGCTCATTAACATCTATG 868

Qy	856	GAGATACCCCTTACACCTGGCATGCTACAAATGGCAAAATTCGAAGTGC	915
Db	869	GTGACACTCCTTTCGACCTGGCATGCTACAAATGGAAATTTTGAAGTGC	928
Qy	916	TCCAAATATCAGGACAGAAAGTCTGACTAAGGAAACACATCTTCAGTGA	975
Db	929	TCAGGTACAGGAACATGAAGTCTGACTAAGGAAACATCTTCAGGAGAG	988
Qy	976	ATAGTGTCTTACCTATGGCAAGACATTCACCTAGTCAAAATCTCTTCT	1035
Db	989	ACAGTGTCTTACCTATGGCAAGAACATTCACCTGTCAAATTTCTCTTGA	1048
Qy	1036	TCTATAAACATCAACCAAGGAAGGATGGGCACACTGGATTAACACTCTG	1095
Db	1049	CTGTGAACATTAACCCAGGAGAGATGGGCACACAGGATTCACCTGCTG	1108
Qy	1096	ACGGTCACATTCGGCTGGTTCAGTCTTACTCGGATAATGGAGCTGATAT	1155
Db	1109	ACGGCCATATCGGCTGGTTCAGTCTTACTTGTATGTTGGTGCAGATAT	1168
Qy	1156	CTTGTGATCCAGCAGGTCTAGTGGTGAATAAGATGAGCAGACATGTTTG	1215
Db	1169	CTTGTGATCCAGCAGGTCTAGTGGTGAATAAGATGAGCAGACATGTTTG	1228
Qy	1216	ATCAAAAGGCATGTCGCATTTGCACACTCTCTGAAGCATTTAAGAGACA	1275
Db	1229	ACGAAAGGACATGATGCCATTTGTTACACTCTGAGCACCTACAGAGACC	1288
Qy	1276	AATGCCCTTGAATGAATATCTCAGCCTGGAGAGATGGCTCTATGTCTGT	1335
Db	1289	AGCTGCCATGTAAAGGAATATTCACGCCTGGAGAGATGGCTCTATGTCT	1348
Qy	1336	CACCTCTGGGGAAGATTAAAGCATGACAAAAGAGAGGCAGATATCTCCT	1395
Db	1349	CCCCCTTGGGCAAGATTAAAGCATGACAAAAGAGAGGCAGATATCTCCT	1408
Qy	1396	CTGGATTCCTTCATATTTCCATTTCTCAGCTCTCAGAAATTGAGTTCAT	1455
Db	1409	CTGAACCTACCCTCCGCTTCCATCTCCAATCTCCGAAATCGAGTTCAC	1468
Qy	1456	GCTCAGGTTCTTTTGGGAAAGTATATAAGGACGATCGAGAAATAAATAG	1515
Db	1469	GCTCGGGTTCCTTTTGGGAAAGTCTATAAGGGCGATGCAAGAAATAAAT	1528
Qy	1516	AACGTTATCGAGCAATACCTACTGTCTCAAAGTCAGATGGATATGTTTT	1575
Db	1529	AACGATACCGAGCAACAACCTACTGCTCCAAGTCAGATGTTTTCCGAG	1588
Qy	1576	TGTCATTTCTCTGCCAGCTCAATCATCTCTGGTGAATTTCAGTTTGGGT	1635
Db	1589	TGTCATTTCTCTGCCAGCTCAACCAACCTCTGGTGGTTTCAGTTTGGGT	1648
Qy	1636	ATGATCCAGCCAGTTTGGCAATGTCACATCAATACATACAGGGGTTCT	1695
Db	1649	ATGACCCCAAGTCAGTTTGGCAATGTCACATCAATACATACAGGGGTT	1708
Qy	1696	TCCTTTCATGACAGAGAGATTCTTGATTTGCAGTCAAAATTAATATTC	1755
Db	1709	TGCTTCATGAACAGAGAAATCTTTGACTTCAGTCTAAATTAATCATTC	1768
Qy	1756	TTGCCAAGGCATCGAGTTCCTTCAAACTCGACACGCAATTAATCATCG	1815
Db	1769	TTGCCAAGGCATCGAGTTCCTGACAGTTTGACCCAGCCAATCATACCG	1828
Qy	1816	ACAGTCACAATATTTCTTCTATGAGGATGGGCATGCTGTGTGGCAGAT	1875
Db	1829	ACAGCCAATATTTCTCTCTATGAGGATGGCCATGCTGTGTGGCAGAT	1888
Qy	1876	CAAGATTTCTACAGTCTCTGGATGAAGACAAATGACAAAACACCTGGG	1935
Db	1889	CAAGATTTCTGACGTCCTCTGGATGAAGACAAATGACAAAACACCTGG	1948
Qy	1936	GGATGGCTCTCTGAGTGTTCACGCAATGTCAGTCTGGTACACATCAAG	1995

Db	1949	GGATGGCCCTGAGGTGTTACAGAGTGCAGAGATACACATCAAGGCTGATGCTTCA	2008
Qy	1996	GCTATGCTCTGCTGCTGGGAAATTTCTCACTGGCGAAATTCATTCGCTCATCTCAAGC	2055
Db	2009	GTTACTCCTGCTGCTGGGAGTCTCTCACTGGAGAAATTCATTCGCTCATCTCAAGC	2068
Qy	2056	CAGCGGCTGGGAGCAGACATATGGGTTACACACATCAGACCTCCCATTTGGCTATCCA	2115
Db	2069	CAGCGGCTGGCAGCAGATATGGGTTATCACCATCAGACCGCCCATCGGCTATCCA	2128
Qy	2116	TTCCCAAGCCATATCATCTCTGCTGATACGAGGGTGAACGGATGCTTGAAGGAGAC	2175
Db	2129	TCCCAAGCCATCATCTCTGCTGATACGAGGGTGAACGGATGCTTGAAGGAGAC	2188
Qy	2176	CCGAATTTCTGAAGTTCTATGAAGTTAGAGAGTGTCTCTGCAACATTTGAGCTGATGT	2235
Db	2189	CAGAGTTCTCTGAAGTCTTATGAACACTGGAGAGTGCCTATGCAATGTGGAGCTCATGT	2248
Qy	2236	CTCCTGCATCAAGTAAACAGAGTGGTCTCTCTCACCTTCTTCTTCTGATTGCTCTG	2295
Db	2249	CTCCAGCATCAAGTAAACAGAGTGGTCTCTCTCACCTTCTTCTTCTGATTGCTCTG	2308
Qy	2296	TGAACCGGGAGGAGCTTGGCGGAGTCTATGTGGCAGCATTAAGAAGTCGTTTCGAATTGG	2355
Db	2309	TGAGCGGGAGGGCTTGGCGGAGCAGCTGGCAGCTTACGAGCGGTTTGAAGTTGG	2368
Qy	2356	AATATGCTTAAATCAAGTCTTATGCTGCTTTGTCGCAAGTCTGCGCAATATTCCT	2415
Db	2369	AGTATGCTTAAATCAAGTCTTATGCTGCTTTGTCGCAAGTCTGCGCAATATTCCT	2428
Qy	2416	CTCAAGTCTGCTTTGAGGAGATGAAGAAAGTCTTCAATACACACCCATTCGCAAAAT	2475
Db	2429	ATCCGGGCTGCTTTGAGGAGATGAAGAAAGTCTTCAATACACACCCATTCGCAAAAT	2488
Qy	2476	ATGGCTATGATTCGATCCCATGAGCTCAATGCAATTTTCAATTTGTCGCGAAATGATGCA	2535
Db	2489	ACGGCTATGCTGCTGATCCCATGAGCTGAGCAGCTTCACTCCCGCAAGAGCAGACGA	2548
Qy	2536	GCTTTGAGCAGCAGCTGACAGCATTCGGGCTTACTCAAGGAGAGTTTTTCCCGCAA	2595
Db	2549	ACTTTGAGCAGCAGCTGACAGCATTCGGGCTTACTCAAGGAGAGTTTTTCCCGCAA	2607
Qy	2596	CTGACAGCAACGATTTCCACCAAGCAGCTGGCTTCCAACTATAAATTTTACTCTCAA	2655
Db	2608	CTGACAGCAGTATTTTACCCTATGCGAGCTTGCTTCCAAATTAAGCGGCTGCGCTCTG	2667
Qy	2656	AGGTCCTTAAATTTGGGCTGTTTACTTGTCTTATTTAAATTTCCCACTATTAGCAGG	2715
Db	2668	AGGT-TTCTTCAATTCGCTTGTCTTATTTAAGCTCGTTTAAATTTCCCTTCTACAGACAG	2726
Qy	2716	CTTTGGATTTGCTTGAAGGATTAATATGCAAAAGAACCAAGACAGATCTATATGAAGA	2775
Db	2727	GCTTTGACTCATGCC--AAGCCTGAAGTGTCAAAGAGCAGATACAGAAATGTCATGAGGA	2784
Qy	2776	ATTGTTTTTAATTTGTTAAATTAATAAATAATTTAGATCGTTACTTGGAAATGGAGCCTA	2835
Db	2785	ATTGTTCTTAGTTTGTATATTTAAAGCCCTTAATTCCTGGGCTGGGGTTCAAA---TC	2840
Qy	2836	AGTCTGTGTGACAGATAAATATTTGTTTCTCTGGGCTGAATATGATGATCTGTGTT	2895
Db	2841	TGTTAGATAGCTGGGTGACCCCTTATGTTTGTAGACCAACCTGTGTGGCTGTGTT	2900
Qy	2896	TGACAG-CTATGGGTTATTTCTTAGAACATTTGTTTCTTTTCTCATTATGTTACT	2954
Db	2901	TGAGGCTCTCTGTTGGGTTTCTTAAATAAACAAGCTGGCTGATTTATCTCTGTTGCTTT	2960
Qy	2955	TCTAGTGTTCACCTCTGTGATTAAGATTTCTTTGTTGAATAGAAAAAATAAAAAA	3014
Db	2961	---GTTGTTACTTCTGTGATTAAGTCTCTTCGGTGATCTAGAAAAAATAAAAAA	3016
Qy	3015	AGGCGGCGC 3022	

Db

3017

GGGCGGCC 3024

RESULT 13

AY303692

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

JOURNAL

FEATURES

source

DBS

16

CGGCCCTGGAGAGGAGAACTTATAAATGGAATTAATAATCTAGACCAACCC

75

12

CAGCAGGAGAAAGCAAGACTTTTAAATGGGAATTAACAATCCAGCAACAC

71

76

AAACTGTGCTGATGAATGGAAGAAAAAGTCAAGTGAATCATATGTTTACAAATAGAAA

135

72

AGACTTGTCTGATGAATGGAAGAAAGTCTTAGTGAATCTTACGCTATTATCATAGAAA

131

136

GATTAGAGATGACCTCAGTCAAGAAAAAGCACTGACAGAACTTAAGAAATATATTTG

195

132

GGCTGGAGGATTAACCTGAGATCAAGAAAAATGAATTTCAAGAACTTAGGCATCTTTG

191

196

GCTCTGATGAAGCCCTTCAGTAAAGTCAATTTAAATTTACCGACTGAAATGGGCTGTC

255

192

GCTCTGATGAAGCCCTTCAGTGAAGTCAAGTTTAAATTTACCGCAGAGCGTGGCTGCCC

251

256

TACTTCATTATGTTGCTTTGAGGAGCAAGAAATACATATTCGAATCTTATGTTGA

315

Db 252 TGCTACACCTCTGCTGCTCTGTGGCGCAACAGTCACATATCCGTCCTTATGTTAA 311  
Qy 316 AAGGCTCCGCCATCTCGACTGACAGAAATGATTTACAGCTTGCATTTAGCAGTTT 375  
Db 312 AAGGCTCCGTCATCCAGACTGACGAAATGGTTTCCAGCTCTGCACCTGGCCGTTT 371  
Qy 376 ACAAGGTAATGACAGATTTGATCACTTCTCTGCTTACAGTGAGTGATATACACAGG 435  
Db 372 ACAGGACAGCCCGAATTTATCACTTCACTGTTGACAGCGAGCAGATGTTACAGAG 431  
Qy 436 TTGATACGTTGGGCTCACTGCCCCATATGCTACATAGTGCCACCTAGAGGCTG 495  
Db 432 TGGATACGTTGGGCTCACAGCCCTCCACATAGCTGCAATAGCTGACACCCAGAGGCTG 491  
Qy 496 CTGATGCTGTTGCAACATGAGCTAATCTCAATATTCAGATGAGATTTTTCACCTC 555  
Db 492 CAGAGTGTCTTACAACATGGGCGCAATGTGAATGTTCAAGATGCGTCTTCTCACCC 551  
Qy 556 CATTGCATATTGACGCTACTATGGACATGAACAGGTAACTCGCCCTCTTTTGAATTTG 615  
Db 552 CACTGCACATTTGACGCTACTATGGGCAAGCAGCAGGTAAACAGTGTCTTTTGAAGTTG 611  
Qy 616 GTGCTGATGTAATGTAAGTGTGAAGTTGAGATAGACCCCTCCACCTAGCATCTGCAA 675  
Db 612 GTGCTGATGTAATGTAAGCGGTGAAGTTGGGACAGGCTCTGCACTGGGCTCTGCAA 671  
Qy 676 AAGGATTTCTTGAATATTGCAAACTCTTGATGGAAGACCAAGCAGATGTGAATG 735  
Db 672 AGGCTTTCTCAACTTGTGAACCTCTGCTAGAGAGGGAGCAAGCAGATGTGAAGC 731  
Qy 736 CTCAGATATGAGACCATGTCCCACTCCATTTCTGTTCTCGATTTGACACCAATGATA 795  
Db 732 CTCAGGCAATGAGACCATGCTCCCTCTGCACTTCTGTTCTCGATTTGACACCAATA 791  
Qy 796 TAGTTAAGTATCTGCTCAAGTGATTTGGAAGTTCAGCTCATGTGTTGTAATCTATG 855  
Db 792 TAGTGACTACCTGCTCCAGATGACTTAGAGTCCAGCTCAGTCAATACATCTATG 851  
Qy 856 GAGATACCCCTTACACCTGGCATGTACAATGGCAAAATTTGAAGTTGCCAAGGAAATCA 915  
Db 852 GTGACACTCTTTGCACCTGGCATGTACAATGGAAATTTGAAGTTGCCAAGGAAATG 911  
Qy 916 TCCAAATACAGGACAGAAAGTCTGACTAAGAAACATCTTCAGTGAACAGCTTTTC 975  
Db 912 TCAGGTTAACAGGAACATGAAGTCTGACTAAGGAAACATCTTCAGCGAGACAGCTTTTC 971  
Qy 976 ATAGTGTCTTGACTATGGCAAGACATTTGACCTAGTCAAATTTCTTGTGATCAGAAATG 1035  
Db 972 ACAGTGTCTTGACTATGGCAAGAACATTTGACCTGTGCAATTTCTTGTGATCAGAAATG 1031  
Qy 1036 TCATAACATCAACCAAGGAGGATGGGCACACTGGATTAACCTCTGCTTGTCTACC 1095  
Db 1032 CTGTGAACATTAACCAAGGAGAGATGGGCACACAGGATTTGCACTCTGCTTGTCTACC 1091  
Qy 1096 ACAGTCACTTCCGCTGTTTCACTTCTTACTGATTAATGGAGCTGATGATCTAGTGG 1155  
Db 1092 ACAGGCAATATCCGCTGTTTCACTTCTTACTGATTAATGGTCAATATGAAATCTTGTG 1151  
Qy 1156 CTGTGATCCAGCAGCTTAGTGGTGAAGAAAGATGAGCAGACATGTTTGTGTTGGGCTT 1215  
Db 1152 CTGTGATCCAGCAGCTTAGTGGTGAAGAAAGATGAGCAGACATGTTTGTGTTGGGCTT 1211  
Qy 1216 ATGAAGAGGGATGATGCAATTTGACACTCTGAGCATTTAAGAGACCAAGATG 1275  
Db 1212 ACAGAAAGGACATGATGCAATTTGTTACACTCTCTGAAGCACTACAAGAGACCCAGGAG 1271  
Qy 1276 AATTGCCCTGTAATGATATTTCTCAGCTGGAGGATGGCTCTATGCTGCTGTTCCAT 1335  
Db 1272 AGCTGCCATGTAACGAATATTTCCAGCTTGGAGGATGGCTCTCTATGTTCTGTTCTCT 1331  
Qy 1336 CACCTTGGGGAAGATTAAGCATGACAAAAGAGAGGACAGATTTCTCTCTTAAGAG 1395

Db 1332 CCCCTTGGGCAAGATTAAAGCATGACAAAAGAGAGGAGATGTTCTCTCCTCCTGAGG 1391  
Qy 1396 CTGATTTGCTTCCATTTCCATCTTCAGCTCTCAGAAATTTGATTTCCATGATATTTG 1455  
Db 1392 CTGAATACCTCCGCTTCCATCTCCAACTCTCCGAAATTCAGATTTCCAGAGATTTATCG 1451  
Qy 1456 GCTCAGGTTCTTTTGGAAAGTATATAAAGGAGATGACAGAAATAAATAGTGGCTATAA 1515  
Db 1452 GCTCGGTTCTTTTGGAAAGTCTATAAAGGCGATGACAGAAATAAATAGTGGCGATCA 1511  
Qy 1516 AAGTTTATCAGGCAATPACTACTGCTCCAAAGTCAGATGTGGATATGTTTTCCGAGAG 1575  
Db 1512 AACGATACCGAGCAACACCTACTGCTCCAAAGTCAGACGCTGGATATGTTTTCCGAGAG 1571  
Qy 1576 TGTCATTTCTCCGAGCTCAATCATCTCCGCTGATTTAGTTTGGGCTGCTGCTTGA 1635  
Db 1572 TGTCATTTCTCCGAGCTCAACACCTCCGCTGGTGGTTTCTGATTTGGGCTGCTGCTGG 1631  
Qy 1636 ATGATCCAGCCAGTTTGCATTTGCTCACTCAATACATATCAGGGGTTCTCTGTTCTCCC 1695  
Db 1632 ATGACCCAGTCAGTTTGCCATTTGCTCACTCAGTACATTTTCCAGAGGCTCCCTGTTCTCCC 1691  
Qy 1696 TCCTTATGAGCAGAGAGGATTTCTGATTTGAGTCTCAATTAATTTATTTGAGTAGATG 1755  
Db 1692 TGCTTCATGAACAGAGAGAAATTTCTGACTTGCAGTCTAAATTAATCATTTGGGAGAG 1751  
Qy 1756 TTGCCAAGCATGTGAGTACTTCAACAACTGACACAGCAATTTATACATCTGACTTGA 1815  
Db 1752 TTGCCAAGGCTGAGTACTTGCACAGCTTGACCCAGCAATCATACACCCGACCTGA 1811  
Qy 1816 ACAGTCAAAATTTCTTCTATGAGATGGGATGCTGTGTGGCAGATTTTGGAGAT 1875  
Db 1812 ACAGCCACAATTTCTGCTCTATGAGATGGCCATGCTGTGTGGCAGATTTTGGAGAT 1871  
Qy 1876 CAAGATTTCTACAGTCTCTGAGTGAAGCAACATGACAAAACAACTGGGACCTCCGTT 1935  
Db 1872 CAAGATTTCTGAGTCCCTGGATGAAGCAACATGACAAAAGCAGCAGGAACTCCGCT 1931  
Qy 1936 GGATGCTCCTGAGGTGTTTCAAGCTGCACTCGGTACACCATCAAGAGCAGATGTCTTCA 1995  
Db 1932 GGATGCTCCTGAGGTGTTTCAAGCTGCAAGATACACCATCAAGGCTGATGTCTTCA 1991  
Qy 1996 GCTATGCTCTGCTGTGTGGGAAATTTCTCACTGGGAAATTTCCATTCGCTCATCTCAAG 2055  
Db 1992 GTTACTCCCTGTCTGTGTGGGAGTCTCTCACTGGGAAATTTCCATTCGCTCATCTCAAG 2051  
Qy 2056 CAGGCTGCTGGCAGCAGACATGGCTTTACCAACATCAGACCTCCCATTTGGCTATTCA 2115  
Db 2052 CAGGCTGCTGGCAGCAGACATGGGCTATCAACATCAGACCGCCCATCGGCTATTCA 2111  
Qy 2116 TTCCCAAGCCCATATCATCTGCTGATACAGGAGTGGAAACGATGCTCTGAAGGAGAC 2175  
Db 2112 TCCCAAGCCCATCTCATCCCTGCTGATACGGGCTGGAAATGCATGCTCTGAAGGAGAC 2171  
Qy 2176 CCAATTTCTGAAGTTGTCAGATTAAGAGTGTCTCTGCAACATTTGAGCTGATGT 2235  
Db 2172 CAGATTTCTGAAGTCTGTTAGCAAACTGGAGAGTGCCTATGCAATTTGGAGCTCATGT 2231  
Qy 2236 CTCTGCATCAAGTAAACAGCAGTGGTCTCTCTCACTTCTTCTTCTTCTGATTTCCCTGG 2295  
Db 2232 CTCCAGCATCAAGTAAACAGCAGTGGTCTCTCTCACTTCTTCTTCTTCTTCCGATTTCCCTGC 2291  
Qy 2296 TGAACCGGAGGACCTGGCCGAGTCTATGGCAGCATTAAGAGTGTGTTTCCGATTTGG 2355  
Db 2292 TGAGCCGGGAGGCTGGCCGAGCAGTGGGAGCTTACGGAGGCTTTTGGATTTGG 2351  
Qy 2356 AATATGCTCTAAATGCAAGTCTCTATGCTTTTCTCCAAAGTCTCGCAATATTTCTCT 2415  
Db 2352 AGTATGCCCTAATGCAAGTCTCTATGCTGGTGTGCCAAAGTGTGGAACACACTCTA 2411  
Qy 2416 CTCAGGCTGTCTTTTGGAGGAGATGAAGAAAGTCTTCAATACACACCCATTTGACAAAT 2475  
Db 2412 ATCCGGGCTGTCTTTTGGAGGAGATGAATGAGGAGCAGCTTATTTCACTGTTGACAAAT 2471

source	1..2505	/organism="Rattus norvegicus"	
		/mol_type="genomic DNA"	
		/db_xref="taxon:10116"	
ORIGIN			
Query Match	62.6%;	Score 1892.2;	DB 6; Length 2505;
Best Local Similarity	84.7%;	Pred. No. 0;	
Matches 2122;	Conservative	0; Mismatches 383;	Indels 0; Gaps 0;
QY	48	ATGGGAATTTAATAATCTAGACCAACCCAAACTTGTACTGTGATGAATGGAAGAAAAATGC	107
Db	1	ATGGGGAATTAACAATCCAGACCAACACAGACTTGTGTTCTGTGATGAATGGAAGAAGAAAT	60
QY	108	AGTGAATCATATGTTATCAATATAGAAAGATTAGAAGATGACCTGCAGATCAAGCAAAAA	167
Db	61	AGTGAATCTTAGCGCTATTATCATAGAAAGGCTGGAGGATAACCTGCAGATCAAGAAAAAT	120
QY	168	GAACGTACAGAACTAAGGAATATATTTGGCTCTGATGAAGCCCTTCAGTAAAGTCAATTTA	227
Db	121	GAATTCAGAACTTAAGGCACATCTTTGGCTCTGATGAAGCCCTTCAGTGAAGTCAATTTA	180
QY	228	AATTACCGCACTGAAAATGGGCTGTCTCTACTTTCATTTATGTTGCAATTTGTGGAGCAAG	287
Db	181	AATTACCGCACAGAGCGTGGCCGTGCTCTGTCTACACCTCTCTGTGCTGTGGCGGCAAC	240
QY	288	AAATCACATATTGCAACTCTTATGTTGAAAGGGCTCCGCCCATCTCGACTGACAAAGAAAT	347
Db	241	AAGTCACATATCGTGGCCCTTATGTTTAAAGGGCTCCGTCCATCCAGACTGACGAGAAAT	300
QY	348	GGATTTACAGCCTTTGCAATTTAGCAGTTTACAAGGATAATGCAAGATTGATCACTTCTCTG	407
Db	301	GGGTTTCCAGCTCTGCACCTGGCCGTTTACAAGGACAGCCCGGAACTTATCACTTCACTG	360
QY	408	CTTCAGAGTGGAGTGAATATACAGCAGGTTGGATACGGTGGCTCACTGCCCTCCATATT	467
Db	361	TTGCACAGCGGAGCAGATGTTTCAGCAAGTGGGATACGGTGGCTTCACGCCCTCCACATA	420
QY	468	GCTACAATAGCTGGCCACCTAGAGGCTGTGATGTCGTGTTGCAACATGGAGCTAATGTC	527
Db	421	GCTGCAATAGCTGGACACCCAGAGGCTGCAGAAAGTGTCTACAACTATGGGGCCAATGTG	480
QY	528	AATATTCAAGATGCAAGTTTTTTTCACTCATTGCAATATTGACGGCTACTATGGACATGAA	587
Db	481	AATGTTCAAGATGCCGCTTCTTTCACCCCACTGCACATTCGAGCTTACTATGGGCACGAG	540
QY	588	CAGGTAACCTCGCCTCTTTTGAATTTGGTGTGATGATAATGTAAGTGGTGAAGTTGGA	647
Db	541	CAGGTAACCAAGTGTCTCTTTTGAAGTTGGTGTGATGTCATGTAAGCGGTGAAGTTGGG	600
QY	648	GATAGACCCCTCCACCTAGCATCTGCAAAAGGATTCTTGAATATTGCAAAAACCTCTTGATG	707
Db	601	GACAGGCTCTGCACCTGGCTCTGCAAAAGGGCTTCTTCAACATTTGTGAATCTCTGGTA	660
QY	708	GAAAGAGGCGACAAAGCAGATGTGAATGCTCAAGATAATGAAGACCATGTCCCACTCCAT	767
Db	661	GAAAGAGGCGACAAAGCAGATGTGAACGCTCAGGACAAATGAAGACCAACGTCCTCTGAC	720
QY	768	TTCTGTTCTCGAATTTGGACACCATGATATAGTTAAGTATCTGCTGCAAAAGTGAATTTGAA	827
Db	721	TTCTGTTCTCGAATTTGGACACCAACAATATAGTAGGCTTACCTGCTCCAGAGTGAATTTAG	780
QY	828	GTTCAACCTCATGTTGTTAATATCTATGGAGATACCCCTTACACCTGGCATGTCTACAAAT	887
Db	781	GTCAGGCTCAGTCATTAACTATGTTGGTGACACTCTCTTTGACCTCTGGCATGTCTACAAAT	840
QY	888	GGCAAAATTTGAAGTTGCCAAGGAATCATCCAAATATCAGGACAGAAAGTCTGACTAAG	947
Db	841	GGAAATTTTGAAGTTGCCAAGGAATTTGTCCAGGTAACAGGAATGAAAGTCTGACTAAG	900
QY	948	GAAAAACATCTTTCAGTGAACAGCTTTTTCATAGTGTGTTTACCTATGCGAAGCATTGAC	1007
Db	901	GAAAAACATCTTTCAGCGAGACAGCTTTTTCAGAGTGTGTTTACCTATGCGAAGAACATTGAC	960

```
QY 1008 CTAGTCAAAATTTCTTTGATCAGAAATGTCTATAAACATCAACACCAAGGAAGGATGGG 1067
Db 961 CTGGTCAAAATTTCTTTGATCAGAAATGTCTATAAACATCAACACCGAGGAGAGATGGG 1020
QY 1068 CACACTGGAATTTACACTCTGCTTGTACACAGGTTCACATTTGGCTGGTTCAGTTCCTACTG 1127
Db 1021 CACACAGGAATTTGCACTCTGCTTGTCTACACAGGCTATATCCGCTCTGCTTCTGCTTCTACTT 1080
QY 1128 GATAATGAGCTGATATGAATTTAGTGGCTTGTGATCCCGAGCAGGCTAGTGGTGAAAAA 1187
Db 1081 GATAATGCTGAGATATGAATTTGTGCTGTGTGATCCCGAGCAGGCTAGTGGTGAAAAA 1140
QY 1188 GATGACGACATATTTGATGTGGCTTTATGAAGGAGGATGATGCCAATGTGCACTC 1247
Db 1141 GATGACGACATATTTGATGTGGCTTTATGAGAAAGGACATGATGCCAATGTGTTACACTC 1200
QY 1248 CTGAAGCATTTAAGAGACCAACAGATGAATTTGCCCTGTAAATTAATTTCTCAGCCTGGA 1307
Db 1201 CTGAAGCACTACAAGACACCCAGAGAGCTGCCATGTAAACGATATTTCCCGAGCCTGGA 1260
QY 1308 GGAGATGCTCCTATGTGTCTGTTCCATCACTCTTGGGGAAGATTAAGAGCATGACAAAA 1367
Db 1261 GGAGATGCTCCTATGTGTCTGTTCCCTTCCCTTGGGCAAGATTAAGAGCATGACAAAA 1320
QY 1368 GAGAAAGCAGATATTTCTCCTTAAGAGCTGGATTCCTTCAATTTCCATCTTCACTC 1427
Db 1321 GAGAAAGCAGATATTTCTCCTGAGGCTGAATCACTACCTCCCGCTTCCATCTTCCAACTC 1380
QY 1428 TCAGAAATGAGTTCCATGAGATTAATGGCTCAGGTTCTTTTGGGAAAGTATATAAGGA 1487
Db 1381 TCCGAAATCGAGTTCCACGAGATTAATCGGCTCGGTTCTTTGGGAAAGTATATAAGGG 1440
QY 1488 CGATGCAAAATTAATTAAGTGGCTATTAAGAGTATGAGCCAAATACCTACTGCTCCAG 1547
Db 1441 CGATGCAAAATTAATTAAGTGGCGATCAACAGATACCGAGCCAAACCTACTGCTCCAG 1500
QY 1548 TCAGATGAGATATTTTGGCAGAGGCTCCATCTCTGCGAGCTCAATCATCCCTGC 1607
Db 1501 TCAGACGTGATATTTTGGCAGAGGCTCCATCTCTGCGAGCTCAACACCCCTGC 1560
QY 1608 GTAATTCAGTTTGGGTGCTTGTGAATGATCCGAGCAGTTGCCATTTGCTACTCAA 1667
Db 1561 GTGGTTTCAGTTTGGGTGCTGCTGATGACCCAGTCAGTTTGCCATTTGCTACTCAG 1620
QY 1668 TACATATCAGGGGTTCTCTGTTCTCCCTCTCATGACGAGAGGATTTCTGATTTG 1727
Db 1621 TACATTTACGAGGCTCCCTGTTCTCCCTCTCATGAAAGAGAAATTTCTGACTTG 1680
QY 1728 CAGTCTAAATTAATTAATGAGTAGATTTGCCAAAGGATGAGTACCTTCACAACTG 1787
Db 1681 CAGTCTAAATTAATTAATGAGTAGATTTGCCAAAGGATGAGTACCTTCACAACTG 1740
QY 1788 ACACAGCAATTTATACATCTGACTTGAACAGTCACAATATTTCTCTATGAGGATGGG 1847
Db 1741 ACCGAGCAATATACACCGGACCTGAAACAGCACAATATTTCTGCTCTATGAGATGGC 1800
QY 1848 CATGCTGTGTGGCAGATTTTGGAGAAATCAAGATTTTCAAGTCTCTGATCAAGACAAAC 1907
Db 1801 CATGCTGTGTGGCAGATTTTGGAGAAATCAAGATTTTCTGAGTCTCCCTGGATGAACAAAC 1860
QY 1908 ATGACAAACAAACCTGGAAACCTCGTTGGATGGCTCTGAGGTGTTTCAGCAGTGCAC 1967
Db 1861 ATGACAAAGCAGCAGGAAACCTCGCTGGATGGCCCTGAGGTGTTTCACACAGTGCACG 1920
QY 1968 CGGTACACCATCAAGCAGATGCTTTCAGTATGCTGCTGCTGCTGGGAAATTTCTCACT 2027
Db 1921 AGATACACCATCAAGGCTGATGCTTCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1980
QY 2028 GGGAAATTCATTCGCTCATCTCAAGCCAGCGGCTGGCGGACGACGATGCTTTACCAAC 2087
Db 1981 GGAGAAATTCATTCGCTCATCTCAAGCCAGCGGCTGGCAGCAGAGATGCTGCTGCTGCTGCT 2040
```

```
QY 2088 CACATCAGACCTCCCATTTGGCTATTCCATTTCCCAAGCCCATATCATCTCTGCTGATACGA 2147
Db 2041 CACATCAGACCGCCCATTCGGCTATTCCATCCCAAGCCCATCTCATCTCTGCTGATACGG 2100
QY 2148 GGTGGAAACGATGCTCTGAAAGAACCCGAAATTTTCTGAAGTTGTCAATGAAGTTAGAA 2207
Db 2101 GGCTGGAAATGATGCTCTGAAAGAACCCGAGAGTTTCTCTGAAGTCGTTAGCAAACTGGAG 2160
QY 2208 CAGTGTCTCTGCAACATTTGAGCTGATGCTCTCTGCTCAAGTAAACAGCAGTGGTCTCTC 2267
Db 2161 GAGTGCCTATGCAATGTGGAGCTCATGTCTCCAGCATCAAGTAAACAGCAGTGGCTCTCTG 2220
QY 2268 TCACCTTCTTCTTCTTCTGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2327
Db 2221 TCACCTTCTTCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2280
QY 2328 GCACATTAAGAGTCTGCTTTCGAATTTGGAATATGCTCTAAATGCAAGGTCCTATGCTGCT 2387
Db 2281 GCACCTTTCAGAGCCGCTTTTGAATTTGGAATATGCTCTAAATGCAAGGTCCTATGCTGCT 2340
QY 2388 TTGTCCCAAAAGTGTGACAAATATTCCTCTCAAGGTCCTGCTTTGGAGGAGATGAAAAA 2447
Db 2341 TGGTCCCAAAAGTGTGGAACACACTCTAATCCGGGCTGCTTTTGGAGGAGATGAATAGG 2400
QY 2448 AGCTTTCATACACACCCATTGACAAATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2507
Db 2401 AGCACCAGATTTCACTGTTGCAAAATACGGCTATGTGTCTGATCCCATGAGCCTGACG 2460
QY 2508 CATTTTCATTTCTTCCGGAATAGTAGCAGCTTTTGAGGACAGCAGC 2552
Db 2461 CACCTTCACCTCCGCAAGACGACAGCAACTTTTGAGGACAGCAAC 2505

RESULT 15
AR269114
LOCUS AR269114 2505 bp DNA linear PAT 10-APR-2003
DEFINITION Sequence 9 from patent US 6500654.
ACCESSION AR269114
VERSION AR269114.1 GI:29699957
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 2505)
AUTHORS Raju, J.
TITLE CARK protein and nucleic acid molecules and uses therefor
JOURNAL Patent: US 6500654-A 9 31-DEC-2002;
FEATURES Location/Qualifiers
source 1..2505
/mol_type="genomic DNA"

ORIGIN
Query Match 62.6%; Score 1892.2; DB 6; Length 2505;
Best Local Similarity 84.7%; Pred. No. 0;
Matches 2122; Conservative 0; Mismatches 383; Indels 0; Gaps 0;

QY 48 ATGGGAAATTTATAAATCTAGACCAACCCAAATCTGCTGATGATGGAAGAAAAGTC 107
Db 1 ATGGGAAATTTACAAATCTCAGACCAACACAGACTTGTCTGATGAATGGAAGAAAAGTT 60
QY 108 AGTGAATCATATGTTTATCAATAGAAAGATTAGAAGATGACCTGCAGATCAAGGAAAAA 167
Db 61 AGTGAATCTTACGCTATTATCATAGAAGGCTGGAGGATAACCTGCAGATCAAGAAAAAT 120
QY 168 GAATCAGACAACTAAGGAATATATTTGGCTCTGATGAAGCCTTCAGTAAAGTCAATTTA 227
Db 121 GAATTTCAAGAACTAAGGCACATCTTTGGCTCTGATGAAGCCTTCAGTAAAGTCAATTTA 180
QY 228 AATTACCGCACTGAAATGGGCTGCTCTTACTTCAATTTATGTTGATTTGAGGAGCAAG 287
Db 181 AATTACCGCAGAGCGTGGCTGTCCCTGCTACACCTCTGCTGCTGCTGCTGCTGCTGCTGCT 240
```

QY	288	AAATCACATATTCGAACTCTTATGTTGAAAGGGCTCGGCCCTCTCGACTGACAGAAAT	347	Db	1321	GAGAGGCAGATGTTCTCTCTCTGAGGGCTGAACCTCCCGCTCCATCTCCAACTC	1380
Db	241	AAGTCACATATCCGTGGCCCTTATGTTAAAGGGCTCGGTCTCCATCCAGACTGACGAGAAAT	300	QY	1428	TCGAAATTCAGTTCATGAGATTATGCTCAGGTTCTTTTGGGAAAGATATATAAGGA	1487
QY	348	GGATTCAGCCTTCGATTTAGCAGTTTACAGGATAATGACAGAAATGATCTCTCTG	407	Db	1381	TCGAAATTCAGTTCACGAGATTATCGGCTCGGTTCTTTTGGGAAAGTCTATAAAGG	1440
Db	301	GGGTTTCAGCTCTGACCTCGCGTTTACAGGACAGCCGGAATCTATCATCTTCACTG	360	QY	1488	CGATGAGAAATAAATAGTGGCTATAAAACGTTATCGAGCCAACTACTGCTCCCAAG	1547
QY	408	CTTCAGTGGAGCTGATATACAGCAGTTGGATACGGTGGCTCTACTGCGCTCCATATT	467	Db	1441	CGATGAGAAATAAATAGTGGCGATCAACAGATACCGAGCCAACTACTGCTCCCAAG	1500
Db	361	TTGCACAGCGAGCAGATGTTTCAGCAAGTGGATACGGTGGCTCTACAGCCCTCCACATA	420	QY	1548	TCAGATGAGATATGTTTTCGAGAGGTGTCATTTCTCTGCGAGCTCAATCATCTCCCTGC	1607
QY	468	GCTACAAATAGCTGGCCACTAGAGCTGCTGATGCTGTGTGCAAGATGAGAGTAATGTC	527	Db	1501	TCAGACGTGATATGTTTTCGCGAGAGGTGTCATTTCTCTGCGAGCTCAACCCACCTGTC	1560
Db	421	GCTGCAATAGCTGGACACCCAGAGGCTGCAGAGTGTCTGTACAAATGGGGCCAATGTG	480	QY	1608	GTAATTCAGTTTCGAGGTGCTGCTGAATGATCCGAGCAGTTTGGCCATTTGCTCACTCAA	1667
QY	528	AAATATCAAGATGCAAGTTTTCCTCACTTCATTTGATATTCAGCGTACTATGACATGAA	587	Db	1561	GTGGTTTCAGTTTCGAGGTGCTGCTGGATGACCCAGTCAGTTTGGCCATTTGCTCACTCAA	1620
Db	481	AAATGTTCAAGATGCGCTCTCTTCACCCCACTGCACATTCAGCCTACTATGGCCACGAG	540	QY	1668	TACATATCAGGGGTTTCTCTGTTTCTCCCTCTTCATGAGCAGAGAGGATTTCTGATTG	1727
QY	588	CAGGTAACCTGCGCTTCTTTTGAATTTGGTCTGATGTAATGTAAGTGTGAGTTGGA	647	Db	1621	TACATTTTCAGAGGCTCCCTGTTTCTCCCTGCTTCATGAAACAGAGAGAAATTTCTGACTG	1680
Db	541	CAGGTAACAGTGTCTTTTGAAGTTTGGTCTGATGTCATGTAAGCGGTGAAGTTGGG	600	QY	1728	CAGTCTAAATTAATTTGAGTAGATGTTGCCAAGGCGATGGAGTACCTTCACAACTG	1787
QY	648	GATAGACCCCTCCACTAGCATCTGCAAAAGGATTTCTTGAATATTCGAAATCTCTTGATG	707	Db	1681	CAGTCTAAATTAATTTGCGGTAGACGTTGCCAAGGCGATGGAGTACCTTCACAACTG	1740
Db	601	GACAGGCTCTGCACCTCGCTCTGCAAGGCTTCTTCAACATTTGTAACCTCTCTGGTA	660	QY	1788	ACACAGCAATTAATATCATGCTGACTTGAACAGTCAATATTTCTTCTCTATGAGGATGGG	1847
QY	708	GAAAGAGCAGCAAGCAGATGTGAATGCTCAAGATTAATGAAGACCATGTCCTCCAT	767	Db	1741	ACCCAGCAATCATACACCGGACCTGGAACGCCACAATATTTCTCTCTATGAGGATGGC	1800
Db	661	GAAAGAGCAGCAAGCAGATGTGAACGCTCAGGACCAATGAAGACCACTGCTCTGCAAC	720	QY	1848	CATCTCTGTTGGCAGATTTTGGAGAAATCAAGATTTCTACAGTCTCTGATGTAAGACAC	1907
QY	768	TTCTGTTCTGATTTGGACACATGATAGTTAAGTATCTGCTGCAAGTGAATTTGGAA	827	Db	1801	CATCTCTGTTGGCAGATTTTGGAGAAATCAAGATTTCTGAGTCTCTGATGTAAGACAC	1860
Db	721	TTCTGTTCTGATTTGGACACACCAATATAGTAGTACCTGCTTCCAGAGTGAATAGAG	780	QY	1908	ATGACAAAACAACTGGGACCTCCGTTGGATGGCTCTGAGGTGTTACGAGTCTCACAGTGCAC	1967
QY	828	GTTCACCTCATGTTGTTTAAATATCTATGAGATACCCCTTTACACCTGGCATCTACAAT	887	Db	1861	ATGACAAAACAACTGGGACCTCCGTTGGATGGCTCTGAGGTGTTACGAGTCTCACAGTGCAC	1920
Db	781	GTCCAGCCTCAGCTCATTAACATCTATGTTGACACTCTCTTTGACCTGGCATGTACAAT	840	QY	1968	CGGTACACCATCAAGCAGATGTTTACGCTATGCTCTGCTGTTGGGAAATTTCTCACT	2027
QY	888	GGCAATTTGAAGTTGCCAAGGAATCATCCAAATATCAGMACAGAAAGTCTGACTAAG	947	Db	1921	AGATACACCATCAAGGCTGATGTTTCAAGTTACTCCCTGTTGTTGGAGTCTCTCACT	1980
Db	841	GGAAATTTTGAAGTTGCCAAGGAATTTGCCAGGTAACAGAACTGAAAGTCTGACTAAG	900	QY	2028	GGCAATTCATTTCCGCTCATCTCAAGCAGCGGCTCGGCGACACACATGCTGTTACCAAC	2087
QY	948	GAAACATCTTCAGTGAAACAGCTTTTCATAGTGTGTTGACCTATGGAAGACATTTGAC	1007	Db	1981	GGAGAAATTCATTTCCGCTCATCTCAAGCAGCGGCTCGGCGACACACATGCTGTTACAC	2040
Db	901	GAAACATCTTCAGTGAAACAGCTTTTCATAGTGTGTTGACCTATGGAAGACATTTGAC	960	QY	2088	CACATCAGACCTCCCATTTGGCTATTCCATTTCCCAAGCCCATATCTCTCTGCTGATACGA	2147
QY	1008	CTAGTCAAAATTTCTTTGATCAGAAATGTCTAAACATCAACACCAAGAGGATGGG	1067	Db	2041	CACATCAGACCGCCCATCGGCTATTCCATCCCAAGCCCATCTCTATCCCTGCTGATACGG	2100
Db	961	CTGGTCAAAATTTCTTTGATCAGAAATGTCTAAACATCAACACCAAGAGGATGGG	1020	QY	2148	GGGTGGAACCGCATGCTCTGAGGAGNAGCCGAAATTTCTGAAAGTCTCATGAGTTAGAA	2207
QY	1068	CACCTGGATTAACCTCTGCTGTACCAAGGCTCAGATTCGCTGGTTCAAGTTCTTACTG	1127	Db	2101	GGGTGGAATGCATGCTCTGAGGACGACAGGTTCTCTGAAAGTCTGTTAGCAAACTGGAG	2160
Db	1021	CACACAGGATTCGATCTGCTGTGCTACCAAGGCTATCCGCTGTTTCAAGTTCTTACTG	1080	QY	2208	CAGTGTCTCTGCAACATTTGAGCTGATGTTCTCTGCAATCAAGTAAACAGCAGTGGTCTCTC	2267
QY	1128	GATAATGAGCTGATAGTAATCTAGTGGCTGTGATCCAGCAGGCTCTAGTGGTGAATAA	1187	Db	2161	CAGTGTCTCTGCAACATTTGAGCTGATGTTCTCTGCAATCAAGTAAACAGCAGTGGTCTCTC	2220
Db	1081	GATAATGAGCTGATAGTAATCTAGTGGCTGTGATCCAGCAGGCTCTAGTGGTGAATAA	1140	QY	2268	TCACCTTCTCTTCTGATGCTGGTGAACCGGGAGGAGCTGGCGGAGTCACTG	2327
QY	1188	GATGAGCAGCATGTTGATGGGCTTATGAAAGGGCATGATGCCATTTGTCACACTC	1247	Db	2221	TCACCTTCTCTTCTCCGATTTGCTGAGCAGGAGGCTGGCGGAGGACACGTTG	2280
Db	1141	GATGAGCAGCATGTTGATGGGCTTATGAAAGGGCATGATGCCATTTGTCACACTC	1200	QY	2328	GCAGCATTAAGAGTCTGTTTCGAATTTGCAATATGCTTAAATCAAGGCTCTATGCTGCT	2387
QY	1248	CTGAACATTAAGAGACCAACAGATGAATTTGCCCTGTAATGAATTTCTCAGCTGGA	1307	Db	2281	GCAGCTTACCGAGCCGCTTTTGTGAGTATGCTTAAATGCAAGTCTCTATGCTGGG	2340
Db	1201	CTGAACATTAAGAGACCAACAGATGAATTTGCCCTGTAATGAATTTCTCAGCTGGA	1260	QY	2388	TTGTCCTCAAAAGTCTGGAACAATTTCTCTCAAGGTTGCTGTTGAGGAGATGAAGA	2447
QY	1308	GAGATGGCTCTTATGTTCTGTTCCATCACCTTGGGAGAGATTAAGCATGACAAA	1367	Db	2341	TGGTCCCAAAAGTGTGGAACAACACTCTAAATCCGCGCTGCTTTGGAGGAGATGAATAG	2400
Db	1261	GGAGATGGCTCTTATGTTCTGTTCCATCACCTTGGGAGAGATTAAGCATGACAAA	1320	QY	2448	AGTCTTCAATACACACCATTTGCAATATGGCTATGATATCCGATGCCATGCTCAATG	2507
QY	1368	GAGAGGCAGATATTTCTCTCTCAAGAGCTGGATGCTTCCATTTTCCATCTTTCAGCTC	1427				

Db 2401 AGCACCCAGTATTCAACTGTGTGACAAATACGGCTATGTGTCTGATCCCATGAGCCTGACG 2460  
QY 2508 CATTTTCATTCTTCCGGAATAGTAGCAGCTTTGAGGACAGCAGC 2552  
Db 2461 CACCTTCACTCCCGCCAAGACGACAGCACTTTGAGGACAGCAAC 2505

Search completed: September 6, 2004, 05:03:12  
Job time : 7790.57 secs

**This Page Blank (uspto)**



Db 61 AGTGAATCATATGTTATCACAAATAGAAGATTAGAAGATGACTGCAGATCAAGGAATAA 120  
QY 121 GAACCTGACAGAACTAAGGAATATATTTGGCTCTGATGAAGCCCTCAGTAAAGTCAATTTA 180  
Db 121 GAACCTGACAGAACTAAGGAATATATTTGGCTCTGATGAAGCCCTCAGTAAAGTCAATTTA 180  
QY 181 AATTACCCGACTGAATAATGGCTCTCTACTTCAATTTATGTTGCAATTTGTCAGGCAAG 240  
Db 181 AATTACCCGACTGAATAATGGCTCTCTACTTCAATTTATGTTGCAATTTGTCAGGCAAG 240  
QY 241 AATACACATATTCGAATCTTTATGTTGAAGGGCTCGGCCCATCTCGACTGACAAGAAAT 300  
Db 241 AATACACATATTCGAATCTTTATGTTGAAGGGCTCGGCCCATCTCGACTGACAAGAAAT 300  
QY 301 GGAATTTACAGCCTTCATTTAGCAGTTTAAGAAGATAATGAGAAATGATCACTTCTCTG 360  
Db 301 GGAATTTACAGCCTTCATTTAGCAGTTTAAGAAGATAATGAGAAATGATCACTTCTCTG 360  
QY 361 CTTACAGTGGAGCTGATATACAGCAGTTTGGATACGGTGGCTCAGTCCCTCCATATT 420  
Db 361 CTTACAGTGGAGCTGATATACAGCAGTTTGGATACGGTGGCTCAGTCCCTCCATATT 420  
QY 421 GGTACAAATAGTGGCCACTAGAGCTGCTGATGTGTTGGAAATGAGCTAATGTC 480  
Db 421 GGTACAAATAGTGGCCACTAGAGCTGCTGATGTGTTGGAAATGAGCTAATGTC 480  
QY 481 AATATTCAGATGCAATTTTTCACCTCATTTGATATGCGAGTACTATGACATGAA 540  
Db 481 AATATTCAGATGCAATTTTTCACCTCATTTGATATGCGAGTACTATGACATGAA 540  
QY 541 CAGGTAAGTCCGCTTCTTTTGAATTTGGTCTGATGTAAATGTAAGTGGTGAAGTTGGA 600  
Db 541 CAGGTAAGTCCGCTTCTTTTGAATTTGGTCTGATGTAAATGTAAGTGGTGAAGTTGGA 600  
QY 601 GATAGACCTCCACTAGCATTCGCAAGGATCTTGAATTTGCAAACTCTTGATG 660  
Db 601 GATAGACCTCCACTAGCATTCGCAAGGATCTTGAATTTGCAAACTCTTGATG 660  
QY 661 GAAGAAGGAGCAAGAGCAGATGTGAATGCTCAAGATAATGAAGACCATGTCCTCCACTCCAT 720  
Db 661 GAAGAAGGAGCAAGAGCAGATGTGAATGCTCAAGATAATGAAGACCATGTCCTCCACTCCAT 720  
QY 721 TTCTGTCTCGATTTGGACACATGATATAGTTAAGTATCTGTGCGAAAGTGAATGGAA 780  
Db 721 TTCTGTCTCGATTTGGACACATGATATAGTTAAGTATCTGTGCGAAAGTGAATGGAA 780  
QY 781 GTTCAACCTCATGTTGTTAATATCTATGGAGATACCCCTTACACCTGGCATGTACAAT 840  
Db 781 GTTCAACCTCATGTTGTTAATATCTATGGAGATACCCCTTACACCTGGCATGTACAAT 840  
QY 841 GGCATAATTTGAAGTTGCCAAGAAATCATCCAAATATCAGAAACAGAAAGTCTGACTAAG 900  
Db 841 GGCATAATTTGAAGTTGCCAAGAAATCATCCAAATATCAGAAACAGAAAGTCTGACTAAG 900  
QY 901 GAAACATCTTCAGTGAACACAGCTTTTCATAGTCTGTGATCTGATGCGAAGCATTGAC 960  
Db 901 GAAACATCTTCAGTGAACACAGCTTTTCATAGTCTGTGATCTGATGCGAAGCATTGAC 960  
QY 961 CTAGTCAAAATTTCTTCTGATCAGAAATGTCATAACATCAACCAAGGAAGGATGGG 1020  
Db 961 CTAGTCAAAATTTCTTCTGATCAGAAATGTCATAACATCAACCAAGGAAGGATGGG 1020  
QY 1021 CACACTGGATATACCTCTGCTTGTACACCGGTCACTTCCCTGGTTCAGTCTTACTG 1080  
Db 1021 CACACTGGATATACCTCTGCTTGTACACCGGTCACTTCCCTGGTTCAGTCTTACTG 1080  
QY 1081 GATAATGGAGCTGATATGAAATCTAGTGGCTTGTGATCCAGCAGGTCTAGTGGTGAAGAA 1140  
Db 1081 GATAATGGAGCTGATATGAAATCTAGTGGCTTGTGATCCAGCAGGTCTAGTGGTGAAGAA 1140  
QY 1141 GATGACACAGATGTTGATGGGCTTATGAAGAAGGGCATGATGCCATGTCACATC 1200

Db 1141 GATGACACAGATGTTTGTGTTGGCTTATGAAAAAGGGCATGATGCCATTTGTCACTC 1200  
QY 1201 CTGAAGCATTTAAGAGACACCAAGATGAATTTGCCCTGTAAATGAATTTCTCAGCCTGGA 1260  
Db 1201 CTGAAGCATTTAAGAGACACCAAGATGAATTTGCCCTGTAAATGAATTTCTCAGCCTGGA 1260  
QY 1261 GGAGATGGCTCCTATGTTCTGTTCCATCACCTTTGGGGAAGATTAAAAAGCATGAACAAA 1320  
Db 1261 GGAGATGGCTCCTATGTTCTGTTCCATCACCTTTGGGGAAGATTAAAAAGCATGAACAAA 1320  
QY 1321 GAGAAGCAGATATTTCTCTCTTAAGAGCTGGATTTGCCCTTCACTTTCCATCTTCAGTCT 1380  
Db 1321 GAGAAGCAGATATTTCTCTCTTAAGAGCTGGATTTGCCCTTCACTTTCCATCTTCAGTCT 1380  
QY 1381 TCAGAAATTCAGTTTCCATGAGATTTATGGCTCAGGTTCTTTTGGGAAAGTATATAAGGA 1440  
Db 1381 TCAGAAATTCAGTTTCCATGAGATTTATGGCTCAGGTTCTTTTGGGAAAGTATATAAGGA 1440  
QY 1441 CGATGCAAAATAAATAGTGGCTATAAAGCTTATCGAGCCAAATACCTACTGCTCCAG 1500  
Db 1441 CGATGCAAAATAAATAGTGGCTATAAAGCTTATCGAGCCAAATACCTACTGCTCCAG 1500  
QY 1501 TCAGATGTGATATGTTTTCGCCGAGAGGTGTCATTTCTGTCAGCTCAATCATCCCTGC 1560  
Db 1501 TCAGATGTGATATGTTTTCGCCGAGAGGTGTCATTTCTGTCAGCTCAATCATCCCTGC 1560  
QY 1561 GTAAATTCAGTTTGTGGTGTCTTGTGAATGATCCAGCCAGTTTGGCAATTTGTCACTCAA 1620  
Db 1561 GTAAATTCAGTTTGTGGTGTCTTGTGAATGATCCAGCCAGTTTGGCAATTTGTCACTCAA 1620  
QY 1621 TACATATCAGGGGTCTCTGTTCTCCCTTCATGAGCAGAAAGGATTTCTGATTTG 1680  
Db 1621 TACATATCAGGGGTCTCTGTTCTCCCTTCATGAGCAGAAAGGATTTCTGATTTG 1680  
QY 1681 CAGTCTAAATTAATTTATTCAGTAGATGTTGCCAAAGGCATGGAGTACCTTCAACCTG 1740  
Db 1681 CAGTCTAAATTAATTTATTCAGTAGATGTTGCCAAAGGCATGGAGTACCTTCAACCTG 1740  
QY 1741 ACACAGCAATATATACATCGTGTGAAACAGTCACAATATTTCTCTCTATCAGGATGG 1800  
Db 1741 ACACAGCAATATATACATCGTGTGAAACAGTCACAATATTTCTCTCTATCAGGATGG 1800  
QY 1801 CATGCTGTGTGGCAGATTTTGGAGATCAAGATTTTACAGTCTCTGATGAAGCAAC 1860  
Db 1801 CATGCTGTGTGGCAGATTTTGGAGATCAAGATTTTACAGTCTCTGATGAAGCAAC 1860  
QY 1861 ATGACAAAACAACTGGGAACTTCCGTTGGATGGCTCTGAGTGTTCAGCAGTGCAT 1920  
Db 1861 ATGACAAAACAACTGGGAACTTCCGTTGGATGGCTCTGAGTGTTCAGCAGTGCAT 1920  
QY 1921 CGGTACACCATCAAGCAGATGCTTTCAGCTATGCTCTGCTGTGGGAAATTTCTCACT 1980  
Db 1921 CGGTACACCATCAAGCAGATGCTTTCAGCTATGCTCTGCTGTGGGAAATTTCTCACT 1980  
QY 1981 GGCAGAAATTCATTCGCTCATCTCAAGCCAGCGCTGGGACAGACATGCTTACCAAC 2040  
Db 1981 GGCAGAAATTCATTCGCTCATCTCAAGCCAGCGCTGGGACAGACATGCTTACCAAC 2040  
QY 2041 CACATCAGACCTCCATTTGGCTATTCCATTTCCAGCCCATATCTCTGCTGATACGA 2100  
Db 2041 CACATCAGACCTCCATTTGGCTATTCCATTTCCAGCCCATATCTCTGCTGATACGA 2100  
QY 2101 GGTGGAAACGCATGCTCTGAGGAAGAACCCGAAATTTCTGAAAGTGTCTCAAGTTAGAA 2160  
Db 2101 GGTGGAAACGCATGCTCTGAGGAAGAACCCGAAATTTCTGAAAGTGTCTCAAGTTAGAA 2160  
QY 2161 GAGTGTCTCGCAACATTTGAGCTGATGCTCTCTGATCAAGTAAACAGAGTGGTCTCTC 2220  
Db 2161 GAGTGTCTCGCAACATTTGAGCTGATGCTCTCTGATCAAGTAAACAGAGTGGTCTCTC 2220  
QY 2221 TCACCTTCTTCTTCTGATTTGCTGGTGAACCGGGGAGGACCTCGCCGAGTCAATG 2280  
Db 2221 TCACCTTCTTCTTCTGATTTGCTGGTGAACCGGGGAGGACCTCGCCGAGTCAATG 2280

```
QY 2281 GCAGCATTAAAGAGTCGTTTCGAATTGGAATATGCTCTMAATGCAAGGTCCTATGCTGCT 2340
Db 2281 GCAGCATTAAAGAGTCGTTTCGAATTGGAATATGCTCTMAATGCAAGGTCCTATGCTGCT 2340
QY 2341 TTGTCCCAAGTGTGACAAATATTCCTCTCAAGGTCGTCTTTGAGGAGATGAAAGA 2400
Db 2341 TTGTCCCAAGTGTGACAAATATTCCTCTCAAGGTCGTCTTTGAGGAGATGAAAGA 2400
QY 2401 AGTCTTCAATACACACCAATGACAAATATGGCTATGATCCGATCCCATGAGCTCAATG 2460
Db 2401 AGTCTTCAATACACACCAATGACAAATATGGCTATGATCCGATCCCATGAGCTCAATG 2460
QY 2461 CATTTTCATTCTTGCCGAATATAGTAGCAGCTTTGAGACAGCAGC 2505
Db 2461 CATTTTCATTCTTGCCGAATATAGTAGCAGCTTTGAGACAGCAGC 2505

RESULT 2
US-10-626-173-3
; Sequence 3, Application US/10626173
; Publication No. US20040110232A1
; GENERAL INFORMATION:
; APPLICANT: Raju, Jeyaseelan
; TITLE OF INVENTION: NOVEL CARP PROTEIN AND NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: MNI-068CP2
; CURRENT APPLICATION NUMBER: US/10/626,173
; CURRENT FILING DATE: 2003-07-24
; PRIOR APPLICATION NUMBER: US/09/947,199A
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: US 09/291,839
; PRIOR FILING DATE: 1999-04-14
; PRIOR APPLICATION NUMBER: US 09/458,457
; PRIOR FILING DATE: 1999-12-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2505
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2505)
US-10-626-173-3

Query Match 100.0%; Score 2505; DB 17; Length 2505;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGAAATTATAAATCTAGACCAACCCAACTTGTAAGTATGATGGAGGAAAGTC 60
Db 1 ATGGGAAATTATAAATCTAGACCAACCCAACTTGTAAGTATGATGGAGGAAAGTC 60
QY 61 AGTGAATCATATGTTATCAATAGAAAGATAGAAAGATGACCTGACATCAAGGAAAA 120
Db 61 AGTGAATCATATGTTATCAATAGAAAGATAGAAAGATGACCTGACATCAAGGAAAA 120
QY 121 GAATGACAGAACTAAGGAATATTTGGCTCTGATGAAGCCCTCAGTAAAGTCAATTA 180
Db 121 GAATGACAGAACTAAGGAATATTTGGCTCTGATGAAGCCCTCAGTAAAGTCAATTA 180
QY 181 AATTACCGCACTGAAATGGGCTGCTCTACTTTCATTTATGTTGATTTGTGGAGGCAAG 240
Db 181 AATTACCGCACTGAAATGGGCTGCTCTACTTTCATTTATGTTGATTTGTGGAGGCAAG 240
QY 241 AAATACATATTGAACTCTTATGTTGAAAGGCTCCGCCATCTCGACTGCAAGAATAAT 300
Db 241 AAATACATATTGAACTCTTATGTTGAAAGGCTCCGCCATCTCGACTGCAAGAATAAT 300
QY 301 GGATTTACGCTTGCATTTAGCAGTTTACAGGATAATGCAAGATTGATCACTTCTCTG 360
Db 301 GGATTTACGCTTGCATTTAGCAGTTTACAGGATAATGCAAGATTGATCACTTCTCTG 360
```

```
Db 301 GGATTTACGCTTGCATTTAGCAGTTTACAGGATAATGCAAGATTGATCACTTCTCTG 360
QY 361 CTTTCACAGTGGAGCTGATATACAGAGGTTGGATACGGTGGCTCACTGCCCTCCATATT 420
Db 361 CTTTCACAGTGGAGCTGATATACAGAGGTTGGATACGGTGGCTCACTGCCCTCCATATT 420
QY 421 GCTCAATATAGCTGGCCACCTAGAGGCTGCTGATGCTCTTTCACACATGAGCTAATGTC 480
Db 421 GCTCAATATAGCTGGCCACCTAGAGGCTGCTGATGCTCTTTCACACATGAGCTAATGTC 480
QY 481 AATATTCAAGATGCAGTCTTTTTCACATCCATGCAATTTGACAGCGTACTATGGACATGA 540
Db 481 AATATTCAAGATGCAGTCTTTTTCACATCCATGCAATTTGACAGCGTACTATGGACATGA 540
QY 541 CAGGTAACCTGCGCTTCTTTTGGAAATTTGGTGTGATGTAATGTAAGTGTGAGTTGGA 600
Db 541 CAGGTAACCTGCGCTTCTTTTGGAAATTTGGTGTGATGTAATGTAAGTGTGAGTTGGA 600
QY 601 GATAGACCCCTCACCTAGCATCTGCAAAAGGATTTTGAATATTGCAAAACTCTTGATG 660
Db 601 GATAGACCCCTCACCTAGCATCTGCAAAAGGATTTTGAATATTGCAAAACTCTTGATG 660
QY 661 GAAGAAGGCGAGCAAGCAGATGTGAATGCTCAAGATAATGAAGACCATGCTCCACTCCAT 720
Db 661 GAAGAAGGCGAGCAAGCAGATGTGAATGCTCAAGATAATGAAGACCATGCTCCACTCCAT 720
QY 721 TTCTGTTCTCGATTTGGACACCATGATATAGTAAATATCTGCTGCAAAAGTATTGGA 780
Db 721 TTCTGTTCTCGATTTGGACACCATGATATAGTAAATATCTGCTGCAAAAGTATTGGA 780
QY 781 GTTCAACCTCATGTTGTTAATATCTATGAGATACCCCTTACACTGCGCATGCTACAA 840
Db 781 GTTCAACCTCATGTTGTTAATATCTATGAGATACCCCTTACACTGCGCATGCTACAA 840
QY 841 GGCAAAATTTGAAGTTGCCAAGGAAATCATCCAAATATCAGGAACAGAAAGTCTGACTAAG 900
Db 841 GGCAAAATTTGAAGTTGCCAAGGAAATCATCCAAATATCAGGAACAGAAAGTCTGACTAAG 900
QY 901 GAAACATCTTCAGTGAACACAGCTTTTCATAGTGTGTTGATGCTTGGCAGAGCAATGAC 960
Db 901 GAAACATCTTCAGTGAACACAGCTTTTCATAGTGTGTTGATGCTTGGCAGAGCAATGAC 960
QY 961 CTAGTCAAAATTTCTCTGATCAGAAATGTATAAATCATCAACCAAGGAGGATGGG 1020
Db 961 CTAGTCAAAATTTCTCTGATCAGAAATGTATAAATCATCAACCAAGGAGGATGGG 1020
QY 1021 CACTGGATTTACATCTGCTGTGCTACACGGTCAATTCGCTGGTTCAGTTCCTTACTG 1080
Db 1021 CACTGGATTTACATCTGCTGTGCTACACGGTCAATTCGCTGGTTCAGTTCCTTACTG 1080
QY 1081 GATAATGGAGCTGATGATGAATCTAGTGGCTTCTGATCCAGCAGGCTAGTGGTGAATA 1140
Db 1081 GATAATGGAGCTGATGATGAATCTAGTGGCTTCTGATCCAGCAGGCTAGTGGTGAATA 1140
QY 1141 GATGAGCAGACATGTTTGTATGTTGGGCTTATGAAAAGGCGCATGATGCCATGTCACATC 1200
Db 1141 GATGAGCAGACATGTTTGTATGTTGGGCTTATGAAAAGGCGCATGATGCCATGTCACATC 1200
QY 1201 CTGAAGCATTTAAGAGACCAAGATGAAATTTGCCCTGTAAATGAATATTTCTCAGGCTGGA 1260
Db 1201 CTGAAGCATTTAAGAGACCAAGATGAAATTTGCCCTGTAAATGAATATTTCTCAGGCTGGA 1260
QY 1261 GGAGATGGCTCTCTATGCTGCTGTTCCATCAGCTTTGGGGAAGATTTAAAGCATGACAAA 1320
Db 1261 GGAGATGGCTCTCTATGCTGCTGTTCCATCAGCTTTGGGGAAGATTTAAAGCATGACAAA 1320
QY 1321 GAGAAAGCAGATATTCTCTCTTAAGAGCTGGAATTCCTTCAATTTCCATTTCCAGTCTC 1380
Db 1321 GAGAAAGCAGATATTCTCTCTTAAGAGCTGGAATTCCTTCAATTTCCATTTCCAGTCTC 1380
QY 1381 TCAGAAATTTGAGTTCCATGAGATTATTTGGCTCAGGTTCTTTTGGGAAAGTATATAAGGA 1440
Db 1381 TCAGAAATTTGAGTTCCATGAGATTATTTGGCTCAGGTTCTTTTGGGAAAGTATATAAGGA 1440
```

1441 QY CGATCGAATAAATAAGTGGCTATATAAGCTTATCGAGCCAAATACCTACTGCTCCAG 1500  
1441 Db |||||  
1501 QY TCAGATGCGATATGTTTTCGAGAGAGTGTCATCTCTGCGAGCTCAATCATCCTCTGC 1560  
1501 Db |||||  
1561 QY GTAAATTCAGTTTGTGGTGTCTGCTTGAATGATCCAGCCAGTGTGCAATGTCATCAA 1620  
1561 Db |||||  
1621 QY TACATATCAGGGGTCTCTGCTTCCCTCCCTCATGAGCAGAGAGGATCTTGATTG 1680  
1621 Db |||||  
1681 QY CAGTCTAAATTAATTTGCAAGTAGATGTTGCCAAGGCGATGGAGTACCTTCACAACCTG 1740  
1681 Db |||||  
1741 QY ACACAGCAATTAATCATCGTGACTTGAACAGTCAATTAATCTCTCATGAGATGG 1800  
1741 Db |||||  
1801 QY CATGCTGTGGTGGCAGATTTTGGAGAAATCAAGATTTCTACAGTCTCTGGATGAACAAC 1860  
1801 Db |||||  
1861 QY ATGACAAAACACCTGGGAACTCCGTTGGATGCTCTGAGTGTTTCACGAGTGCAC 1920  
1861 Db |||||  
1921 QY CGGTACACCATCAAGCAGATGCTTCAAGTATGCTCTGCTGTCTGGGAAATTCAC 1980  
1921 Db |||||  
1981 QY GCGAAATTCATTCGCTCATCTCAAGCCAGCGCTCGGAGCAGACATGGCTTACCAC 2040  
1981 Db |||||  
2041 QY CACATCAGACCTCCCATTTGGCTATTCATTCGAGCCCATATCATCTCTGCTGATACGA 2100  
2041 Db |||||  
2101 QY GGTGGAACGCGATGCTTGAAGAGAGACCGAAATTTCTGAAGTTGTCATGAAGTTAGAA 2160  
2101 Db |||||  
2161 QY GAGTGTCTGCAACATTTGAGTGTCTCTGATCAAGTACAGCAGTGGTCTCTC 2220  
2161 Db |||||  
2221 QY TCACCTTCTTCTTCTTCTGATTTGCTGTGAACCGGGAGACCTGCGCGAGTCAATGTG 2280  
2221 Db |||||  
2281 QY GCAGCATTAAAGAGTCGTTTGAATTTGAATGATGCTCTAAATGCAAGGTCCTATGCTGCT 2340  
2341 QY TTGTCGCAAGTGTGCAATATTCCTCTCAAGGTCGCTTCTTGGAGGAGATGAAGA 2400  
2341 Db |||||  
2401 QY AGTCTTCAATACACACCCATTCACAAATATGCTATGATCCGATCCCATGAGTCAATG 2460  
2401 Db |||||  
2461 QY CATTTTCATTCTTGGCGAAATAGTAGAGCTTTGAGGACAGCAGC 2505  
2461 Db |||||

RESULT 3  
US-09-947-199-1  
; Sequence 1, Application US/09947199  
; Patent No. US20020127684A1  
; GENERAL INFORMATION:  
; APPLICANT: Raju, Jeyaseelan  
; TITLE OF INVENTION: NOVEL CARP PROTEIN AND NUCLEIC ACID MOLECULES AND USES  
; TITLE OF INVENTION: THEREFOR  
; FILE REFERENCE: MN1-068P2  
; CURRENT APPLICATION NUMBER: US/09/947,199  
; CURRENT FILING DATE: 2001-09-05  
; PRIOR APPLICATION NUMBER: 60/111,938  
; PRIOR FILING DATE: 1998-12-11  
; PRIOR APPLICATION NUMBER: 09/291,839  
; PRIOR FILING DATE: 1999-04-14  
; PRIOR APPLICATION NUMBER: 09/458,457  
; PRIOR FILING DATE: 1999-12-10  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 3025  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (48)..(2552)  
US-09-947-199-1  
  
Query Match 100.0%; Score 2505; DB 9; Length 3025;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 ATGGGAATTAATAATCTAGACCAACCCAAACCTTGATGATGAATGGAAGAAAAAGTC 60  
Db 48 ATGGGAATTAATAATCTAGACCAACCCAAACCTTGATGATGAATGGAAGAAAAAGTC 107  
  
QY 61 AGTGAATCATATGTTATCAATAGAAAGTTAGAAAGTACCTGAGATCAAGAAAA 120  
Db 108 AGTGAATCATATGTTATCAATAGAAAGTTAGAAAGTACCTGAGATCAAGAAAA 167  
  
QY 121 GAATGACAGAACTAAGGAATATATTTGGCTCTGATGAAGCCCTTCAGTAAAGTCAATT 180  
Db 168 GAATGACAGAACTAAGGAATATATTTGGCTCTGATGAAGCCCTTCAGTAAAGTCAATT 227  
  
QY 181 AATTACCGCACTGAAATGGGCTGCTCTACTTCTTATGTTGCTATTTGTGGAGCGAG 240  
Db 228 AATTACCGCACTGAAATGGGCTGCTCTACTTCTTATGTTGCTATTTGTGGAGCGAG 267,  
  
QY 241 AATTCACATATTCGAACCTTTATGTTGAAAGGGCTCGGCCATCTCGACTGACAAAGAA 300  
Db 288 AATTCACATATTCGAACCTTTATGTTGAAAGGGCTCGGCCATCTCGACTGACAAAGAA 347  
  
QY 301 GGAATTCAGCCTTGCATTTAGCAGTTTACAGGATTAACAGGATTAATGATGATCTCTCTG 360  
Db 348 GGAATTCAGCCTTGCATTTAGCAGTTTACAGGATTAACAGGATTAATGATGATCTCTCTG 407  
  
QY 361 CTTACAGTGGAGCTGATATACAGAGTTGGATGAGTGGCTGCTACTGCTCCATATTC 420  
Db 408 CTTACAGTGGAGCTGATATACAGAGTTGGATGAGTGGCTGCTACTGCTCCATATTC 467  
  
QY 421 GCTACAATAGCTGCCACCTAGAGGCTGCTGATGCTGTGTGCAACATCGAGCTAATGTC 480  
Db 468 GCTACAATAGCTGCCACCTAGAGGCTGCTGATGCTGTGTGCAACATCGAGCTAATGTC 527  
  
QY 481 AATATTCAGATGAGGTTTTTTTCACTCCATTCATATTCAGCGTACTATGACATGAA 540  
Db 528 AATATTCAGATGAGGTTTTTTTCACTCCATTCATATTCAGCGTACTATGACATGAA 587  
  
QY 541 CAGTAACTCGCCTTCTTTTGAATTTGGTGTGCTGATGTAATGTAAGTGGTGAAGTTGA 600  
Db 588 CAGTAACTCGCCTTCTTTTGAATTTGGTGTGCTGATGTAATGTAAGTGGTGAAGTTGA 647

QY 601 GATAGACCCCTCCACCTAGCATCTGCAAAAGGATTCCTTGAATATTCGAAAACCTCTTGATG 660  
DB 648 GATAGACCCCTCCACCTAGCATCTGCAAAAGGATTCCTTGAATATTCGAAAACCTCTTGATG 707  
QY 661 GAAGAAGCGACGAAGAGATGTAAGTCTCAAGATAAATGAAGACCATGTCCCACTCCAT 720  
DB 708 GAAGAAGCGACGAAGAGATGTAAGTCTCAAGATAAATGAAGACCATGTCCCACTCCAT 767  
QY 721 TTCTGTCTCGATTTGGACACCAATGATATAGTATCTGCTGCAAAAGTGAATTTGGAA 780  
DB 768 TTCTGTCTCGATTTGGACACCAATGATATAGTATCTGCTGCAAAAGTGAATTTGGAA 827  
QY 781 GTTCAACCTCATGTTGTAATATCTATGAGATACCCCTTACACCTGCGCATGTACAT 840  
DB 828 GTTCAACCTCATGTTGTAATATCTATGAGATACCCCTTACACCTGCGCATGTACAT 887  
QY 841 GGCAAAATTTGAAGTTGCCAAGGAATCATCCAAATATCAGGAACAGAAAGTCTGACTAAG 900  
DB 888 GGCAAAATTTGAAGTTGCCAAGGAATCATCCAAATATCAGGAACAGAAAGTCTGACTAAG 947  
QY 901 GAAACACATCTTCAGTGAACAGCTTTTCATAGTGGCTGTACCTATGGCAAGGACATTGAC 960  
DB 948 GAAACACATCTTCAGTGAACAGCTTTTCATAGTGGCTGTACCTATGGCAAGGACATTGAC 1007  
QY 961 CTAGTCAAAATTTCTCTTGATCAGAAATGTCATAAATCAACCAACCAAGGAAGGATGGG 1020  
DB 1008 CTAGTCAAAATTTCTCTTGATCAGAAATGTCATAAATCAACCAACCAAGGAAGGATGGG 1067  
QY 1021 CACACTGGATACACTCTGCTGTGCTACCGGTGACATTCGCGCTGTTGAGTTCATTCTG 1080  
DB 1068 CACACTGGATACACTCTGCTGTGCTACCGGTGACATTCGCGCTGTTGAGTTCATTCTG 1127  
QY 1081 GATAATCGAGCTGATAGTAATAGTGGCTGTGATCCAGCAGGTCTAGTGTGAAAAA 1140  
DB 1128 GATAATCGAGCTGATAGTAATAGTGGCTGTGATCCAGCAGGTCTAGTGTGAAAAA 1187  
QY 1141 GATGAGCAGACATGTTTGATGTGGGCTTATGAAAAGGGGATGATGATGATGATGATG 1200  
DB 1188 GATGAGCAGACATGTTTGATGTGGGCTTATGAAAAGGGGATGATGATGATGATGATG 1247  
QY 1201 CTGAGCATTATAAGACACACAGATGATTCGCCCTGTAAATGATATTCAGCCTGGA 1260  
DB 1248 CTGAGCATTATAAGACACACAGATGATTCGCCCTGTAAATGATATTCAGCCTGGA 1307  
QY 1261 GGAGATGGCTCCTATGCTGTGTTCCATCAACCCCTGGGGAAGATTAAAGCATGACAAA 1320  
DB 1308 GGAGATGGCTCCTATGCTGTGTTCCATCAACCCCTGGGGAAGATTAAAGCATGACAAA 1367  
QY 1321 GAGAAGCAGATATTCCTCCTTAAGCTGAGATTCCTTCATTTCCATTTCCATTTCCAGCTC 1380  
DB 1368 GAGAAGCAGATATTCCTCCTTAAGCTGAGATTCCTTCATTTCCATTTCCATTTCCAGCTC 1427  
QY 1381 TCAGAAATTCAGTTCATGAGATTTATGGCTCAGGTTCTTTTGGGAAAGTATATAAGGA 1440  
DB 1428 TCAGAAATTCAGTTCATGAGATTTATGGCTCAGGTTCTTTTGGGAAAGTATATAAGGA 1487  
QY 1441 CGATGCAAAATAAATAGTGGCTATAAAGCGTTATCGAGCCAAATACCTACTGCTCCAAG 1500  
DB 1488 CGATGCAAAATAAATAGTGGCTATAAAGCGTTATCGAGCCAAATACCTACTGCTCCAAG 1547  
QY 1501 TCAGATGAGATATGTTTCCGAGAGGTGTCATTTCTGCGCAGCTCAATCATCCCTGC 1560  
DB 1548 TCAGATGAGATATGTTTCCGAGAGGTGTCATTTCTGCGCAGCTCAATCATCCCTGC 1607  
QY 1561 GTAATTCAGTTTGGGTGCTGTGTTGAATGATCCAGCCAGTTTGCCATTTGCTACTCAA 1620  
DB 1608 GTAATTCAGTTTGGGTGCTGTGTTGAATGATCCAGCCAGTTTGCCATTTGCTACTCAA 1667  
QY 1621 TACATATCAGGGGTTCTCTGTTCTCCCTCTCATGAGCAGAGGATTTCTGATTTG 1680  
DB 1668 TACATATCAGGGGTTCTCTGTTCTCCCTCTCATGAGCAGAGGATTTCTGATTTG 1727  
QY 1681 CAGTCTAAATTAATATTGCGATAGATGTTGGCCAAAGGATGAGTACCTTCACAACCTG 1740

DB 1728 CAGTCTAAATTAATATTATGAGTAGATGTTGCAAAAGGATGAGTACCTTCACAACCTG 1787  
QY 1741 ACACAGCCAAATTAACATCGTACCTTGAACAGTCACAATATTCTTCTATGAGGATGG 1800  
DB 1788 ACACAGCCAAATTAACATCGTACCTTGAACAGTCACAATATTCTTCTATGAGGATGG 1847  
QY 1801 CATGCTGTGTGGCAGATTTTGGAGAAATCAAGATTTCTACAGTCTCTGGATGAAGCAAC 1860  
DB 1848 CATGCTGTGTGGCAGATTTTGGAGAAATCAAGATTTCTACAGTCTCTGGATGAAGCAAC 1907  
QY 1861 ATGACAAAACAACTGGGAACCTCGTGGATGGCTCCCTGAGGTGTTACGCAAGTGCACT 1920  
DB 1908 ATGACAAAACAACTGGGAACCTCGTGGATGGCTCCCTGAGGTGTTACGCAAGTGCACT 1967  
QY 1921 CGGTACACCATCAAGCAGATGCTTTCAGTATGCTCTGTGTGTGGGAAATTTCTCACT 1980  
DB 1968 CGGTACACCATCAAGCAGATGCTTTCAGTATGCTCTGTGTGTGGGAAATTTCTCACT 2027  
QY 1981 GGGGAATTCATTCCTCATCTCAAGCCAGCGGTCGGCAGCAGACATGCTTACCAC 2040  
DB 2028 GGGGAATTCATTCCTCATCTCAAGCCAGCGGTCGGCAGCAGACATGCTTACCAC 2087  
QY 2041 CACATCAGACCTCCCATTTGGCTATTCCATTCCCAAGCCCATATCATCTCTGCTGATAGA 2100  
DB 2088 CACATCAGACCTCCCATTTGGCTATTCCATTCCCAAGCCCATATCATCTCTGCTGATAGA 2147  
QY 2101 GGGTGGAAACATGCTCTGAAAGAACCCGGAATTTCTGAAGTTGTCAAGAGTTAGAA 2160  
DB 2148 GGGTGGAAACATGCTCTGAAAGAACCCGGAATTTCTGAAGTTGTCAAGAGTTAGAA 2207  
QY 2161 GAGTGTCTCTGCAACATTTGAGCTGATGCTCTCGCATCAAGTAAACAGCAGTGGGTCTCTC 2220  
DB 2208 GAGTGTCTCTGCAACATTTGAGCTGATGCTCTCGCATCAAGTAAACAGCAGTGGGTCTCTC 2267  
QY 2221 TCACCTTCTTCTTCTCTGATGCTGTTGAACCGGGGAGGACCTGGCCGGAGTCAATGG 2280  
DB 2268 TCACCTTCTTCTTCTCTGATGCTGTTGAACCGGGGAGGACCTGGCCGGAGTCAATGG 2327  
QY 2281 GCAGCATTAAGAGTCTGTTTGAATTTGAATATGCTCTAAATGCAAGGTCTTATGCTGT 2340  
DB 2328 GCAGCATTAAGAGTCTGTTTGAATTTGAATATGCTCTAAATGCAAGGTCTTATGCTGT 2387  
QY 2341 TTGTCCAAAGTCTCGCAATATTCCTCTCAAGTCTCTTTTGGAGGAGATGAAAAGA 2400  
DB 2388 TTGTCCAAAGTCTCGCAATATTCCTCTCAAGTCTCTTTTGGAGGAGATGAAAAGA 2447  
QY 2401 AGTCTTCAATACACACCCATGACAAATATGGCTATGATCCGATCCCATGAGCTCAATG 2460  
DB 2448 AGTCTTCAATACACACCCATGACAAATATGGCTATGATCCGATCCCATGAGCTCAATG 2507  
QY 2461 CATTTTCATTTCTTGGCCGAAATAGTAGCAGCTTTTGAGCAGCAGC 2505  
DB 2508 CATTTTCATTTCTTGGCCGAAATAGTAGCAGCTTTTGAGCAGCAGC 2552

## RESULT 4

US-10-626-173-1  
; Sequence 1, Application US/10626173  
; Publication No. US20040110232A1  
; GENERAL INFORMATION:  
; APPLICANT: Raju, Jeyaseelan  
; TITLE OF INVENTION: NOVEL CARX PROTEIN AND NUCLEIC ACID MOLECULES AND USES  
; FILE REFERENCE: MN-068CP2  
; CURRENT FILING DATE: 2003-07-24  
; PRIOR APPLICATION NUMBER: US 60/111,938  
; PRIOR FILING DATE: 1998-12-11  
; PRIOR APPLICATION NUMBER: US 09/291,839  
; PRIOR FILING DATE: 1999-04-14  
; PRIOR APPLICATION NUMBER: US 09/458,457

;/ PRIOR FILING DATE: 1999-12-10  
;/ NUMBER OF SEQ ID NOS: 9  
;/ SOFTWARE: PatentIn Ver. 2.0  
;/ SEQ ID NO 1  
;/ LENGTH: 3025  
;/ TYPE: DNA  
;/ ORGANISM: Homo sapiens  
;/ FEATURE:  
;/ NAME/KEY: CDS  
;/ LOCATION: (48)..(2552)  
US-10-626-173-1

Query Match 100.0%; Score 2505; DB 17; Length 3025;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGAATTTAAATCTAGACCAACCCAACTTGTACTGATGAATGGAAGAAAAGTC 60  
DB 48 ATGGGAATTTAAATCTAGACCAACCCAACTTGTACTGATGAATGGAAGAAAAGTC 107  
QY 61 AGTGAATCATATGTTTATCAATAGAAAGATTAGAAGATGACCTGCGAGATCAAGGAAAAA 120  
DB 108 AGTGAATCATATGTTTATCAATAGAAAGATTAGAAGATGACCTGCGAGATCAAGGAAAAA 167  
QY 121 GAACCTGACAGAACTAAGAAATATTTGGCTCTGATGAAGCCTTCAGTAAAGTCAATTTA 180  
DB 169 GAACCTGACAGAACTAAGAAATATTTGGCTCTGATGAAGCCTTCAGTAAAGTCAATTTA 227  
QY 181 AATTACCGCACTGAAAATGGGCTGTCTACTTCATTATATTTGTCATTTGTGAGGCAAG 240  
DB 228 AATTACCGCACTGAAAATGGGCTGTCTACTTCATTATATTTGTCATTTGTGAGGCAAG 287  
QY 241 AATACACATATTCGAACCTTTATGTTGAAGGGCTCCGCCATCTCGACTGCAAGAAAT 300  
DB 288 AATACACATATTCGAACCTTTATGTTGAAGGGCTCCGCCATCTCGACTGCAAGAAAT 347  
QY 301 GGATTTACAGCCCTTCGCAATTTAGCAGTTTACAGGATAATGCAAGATTTGATCACTTCTCTG 360  
DB 348 GGATTTACAGCCCTTCGCAATTTAGCAGTTTACAGGATAATGCAAGATTTGATCACTTCTCTG 407  
QY 361 CTTTCAGTGGAGCTGATATACAGCAGTTGGATACGGTGGCTCCTCAGTCCCTCCATATT 420  
DB 408 CTTTCAGTGGAGCTGATATACAGCAGTTGGATACGGTGGCTCCTCAGTCCCTCCATATT 467  
QY 421 GCTACAATAGCTGGCCACTAGAGGCTGCTGATGCTGTTGCAACATGAGGCTAATGTC 480  
DB 468 GCTACAATAGCTGGCCACTAGAGGCTGCTGATGCTGTTGCAACATGAGGCTAATGTC 527  
QY 481 AATATTCAGATGCGAGTTTCTTCTCCTCATTTGATATGTCAGGCTACTATGACATGAA 540  
DB 528 AATATTCAGATGCGAGTTTCTTCTCCTCATTTGATATGTCAGGCTACTATGACATGAA 587  
QY 541 CAGGTAACTCGGCTTCTTTGAAATTTGGTCTGATGTTAAATGTAAGTGGTGAAGTTGGA 600  
DB 588 CAGGTAACTCGGCTTCTTTGAAATTTGGTCTGATGTTAAATGTAAGTGGTGAAGTTGGA 647  
QY 601 GATAGACCCCTCCACTAGCATCTGCAAAAGGATTTCTTGAATATGCAAAACCTTTGATG 660  
DB 648 GATAGACCCCTCCACTAGCATCTGCAAAAGGATTTCTTGAATATGCAAAACCTTTGATG 707  
QY 661 GAAAGAGGCAAGACAGATGTAATGCTCAAGATAATGAGACATGTCCTCCACTCCAT 720  
DB 708 GAAAGAGGCAAGACAGATGTAATGCTCAAGATAATGAGACATGTCCTCCACTCCAT 767  
QY 721 TTCTGTTCTCGATTGGACACCATTGATATAGTAAATGATCTGCTGCAAAAGTGAATTTGGA 780  
DB 768 TTCTGTTCTCGATTGGACACCATTGATATAGTAAATGATCTGCTGCAAAAGTGAATTTGGA 827  
QY 781 GTTCAACCTCATGTTGTTAATATCTATGAGATACCCCTTTACACCTGGCATGCTACAAT 840  
DB 828 GTTCAACCTCATGTTGTTAATATCTATGAGATACCCCTTTACACCTGGCATGCTACAAT 887  
QY 841 GGCAAAATTTGAAGTTGCCAAGGAAATCATCCAAATATACAGGAACAGAAAGTCTGACTAAG 900

DB 888 GGCAAAATTTGAAGTTGCCAAGGAAATCATCCAAATATCAGGAACAGAAAGTCTGACTAAG 947  
QY 901 GAAACATCTTTCACTGAAACAGCTTTTTCATAGTCTTGTACCTATGGCAAGCATTGAC 960  
DB 948 GAAACATCTTTCACTGAAACAGCTTTTTCATAGTCTTGTACCTATGGCAAGCATTGAC 1007  
QY 961 CTAGTCAAAATTTCTTCTTGTATGAGATGTCATAAACATCAACCAACCAAGAGGATGGG 1020  
DB 1008 CTAGTCAAAATTTCTTCTTGTATGAGATGTCATAAACATCAACCAACCAAGAGGATGGG 1067  
QY 1021 CACACTGATTTACACTCTGCTTACCAAGGTCACATTCGGCTGTGTTTCTGATTTCTTACTG 1080  
DB 1068 CACACTGATTTACACTCTGCTTACCAAGGTCACATTCGGCTGTGTTTCTGATTTCTTACTG 1127  
QY 1081 GATAATGAGCTGATATGAATCTAGTGGCTTGTGATCCCAAGAGGCTTGTAGTGGTGAAGAA 1140  
DB 1128 GATAATGAGCTGATATGAATCTAGTGGCTTGTGATCCCAAGAGGCTTGTAGTGGTGAAGAA 1187  
QY 1141 GATGAGCAGACATGTTTGTATGCTGGCTTATGAAAAGGCGATGATGCCATTTGTCACTC 1200  
DB 1188 GATGAGCAGACATGTTTGTATGCTGGCTTATGAAAAGGCGATGATGCCATTTGTCACTC 1247  
QY 1201 CTGAAGCATTTAAGAGACCAAGATGAATTTGCCCTGTAAATGAATTTCTCAGCTGGA 1260  
DB 1248 CTGAAGCATTTAAGAGACCAAGATGAATTTGCCCTGTAAATGAATTTCTCAGCTGGA 1307  
QY 1261 GGAGATGGCTCTATGCTGTCTCCATCACCTTGGGAGAGATTAAGAGCATGACAAA 1320  
DB 1308 GGAGATGGCTCTATGCTGTCTCCATCACCTTGGGAGAGATTAAGAGCATGACAAA 1367  
QY 1321 GAGAAGCAGATATTTCTCTCTAAGAGCTGGAATTCCTTCACATTTCCATCTTCCAGCTC 1380  
DB 1368 GAGAAGCAGATATTTCTCTCTAAGAGCTGGAATTCCTTCACATTTCCATCTTCCAGCTC 1427  
QY 1381 TCAGAAATGAGTTCCATGAGATTTTGGCTCAGGTTCTTTTGGGAAAGTATATAAGGA 1440  
DB 1428 TCAGAAATGAGTTCCATGAGATTTTGGCTCAGGTTCTTTTGGGAAAGTATATAAGGA 1487  
QY 1441 CGATGCAAAATTAATAGTGGCTTATAAACGTTATCGAGCCAAATACCTTCTCTCCAG 1500  
DB 1488 CGATGCAAAATTAATAGTGGCTTATAAACGTTATCGAGCCAAATACCTTCTCTCCAG 1547  
QY 1501 TCAGATGAGATATGTTTTCAGAGAGGTTCCATTTCTGCGAGCTCAATCATCTCCCTG 1560  
DB 1548 TCAGATGAGATATGTTTTCAGAGAGGTTCCATTTCTGCGAGCTCAATCATCTCCCTG 1607  
QY 1561 GTRATTCAGTTTGGGCTGCTTGTGATGATCCAGAGGTTTGGCTTCTTCTTCTTCTTCT 1620  
DB 1608 GTRATTCAGTTTGGGCTGCTTGTGATGATCCAGAGGTTTGGCTTCTTCTTCTTCTTCT 1667  
QY 1621 TACATATCAGGGGTTCTCTGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1680  
DB 1668 TACATATCAGGGGTTCTCTGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1727  
QY 1681 CAGCTTAAATTAATTTGAGATGTTGCCAAAGGATGAGTACCTTCCACACCTG 1740  
DB 1728 CAGCTTAAATTAATTTGAGATGTTGCCAAAGGATGAGTACCTTCCACACCTG 1787  
QY 1741 ACAGAGCAATTAATATACATCGTACTTGAACAGTCAATATTTCTTCTTCTTCTTCTTCT 1800  
DB 1788 ACAGAGCAATTAATATACATCGTACTTGAACAGTCAATATTTCTTCTTCTTCTTCTTCT 1847  
QY 1801 CATGCTGTTGGGAGATTTTGGAGATTAAGATTTCTCAGTCTCTGATGATGAAGACAC 1860  
DB 1848 CATGCTGTTGGGAGATTTTGGAGATTAAGATTTCTCAGTCTCTGATGATGAAGACAC 1907  
QY 1861 ATGCAAAAACAACTGGGAACTCCCGTTGGATGGCTCTTCTGAGTGTTCACGAGTGCCT 1920  
DB 1908 ATGCAAAAACAACTGGGAACTCCCGTTGGATGGCTCTTCTGAGTGTTCACGAGTGCCT 1967  
QY 1921 CGGTACACCATCAAGAGATGCTTTCAGCTATGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 1980

Db 1968 CGGTACACCATCAAGAGAGTCTTCAGGTATGCTCTGTGTCTGTGGGAAATTTCTCACT 2027  
Qy 1981 GGGCAAAATTCATTCGCTCATCTCAAGCCAGCGCTGGCGCAGACACATGCTTACCAAC 2040  
Db 2028 GGGCAAAATTCATTCGCTCATCTCAAGCCAGCGCTGGCGCAGACACATGCTTACCAAC 2087  
Qy 2041 CACATCAGACCTCCATTCGCTATTCATTCACAGCCCATATCATCTCTGCTGATACGA 2100  
Db 2088 CACATCAGACCTCCATTCGCTATTCATTCACAGCCCATATCATCTCTGCTGATACGA 2147  
Qy 2101 GGGTGGAAACGATCTCTGAAGGAAGACCCGAATTTCTGAAGTCTCATGAAGTTAGAA 2160  
Db 2148 GGGTGGAAACGATCTCTGAAGGAAGACCCGAATTTCTGAAGTCTCATGAAGTTAGAA 2207  
Qy 2161 GAGTGTCTCTGCAACATGAGCTGATCTCTCTGCAATCAAGTAAACAGAGTGGTCTCTC 2220  
Db 2208 GAGTGTCTCTGCAACATGAGCTGATCTCTCTGCAATCAAGTAAACAGAGTGGTCTCTC 2267  
Qy 2221 TCACCTCTCTCTCTCTCTGATGCTGTAAGCGGGAGGACCTGGCGGAGTCAATG 2280  
Db 2268 TCACCTCTCTCTCTCTCTGATGCTGTAAGCGGGAGGACCTGGCGGAGTCAATG 2327  
Qy 2281 GCAGCATTAAGAAAGTCTGTCGAATTTGGAATATGCTCTAAATGCAAGTCTCTATGCTGT 2340  
Db 2328 GCAGCATTAAGAAAGTCTGTCGAATTTGGAATATGCTCTAAATGCAAGTCTCTATGCTGT 2387  
Qy 2341 TTGTCCAAAGTCTGACAAATATCTCTCAAGTCTGTCTTTGGAGAGATGAAAGA 2400  
Db 2388 TTGTCCAAAGTCTGACAAATATCTCTCAAGTCTGTCTTTGGAGAGATGAAAGA 2447  
Qy 2401 AGTCTTCAATACACACCATGTCGAATATGCTATGCTATGCTATGCTATGCTATGCTCAATG 2460  
Db 2448 AGTCTTCAATACACACCATGTCGAATATGCTATGCTATGCTATGCTATGCTCAATG 2507  
Qy 2461 CATTTTCATCTTGGCCAAATAGTAGAGCTTTTGGAGACAGCAGC 2505  
Db 2508 CATTTTCATCTTGGCCAAATAGTAGAGCTTTTGGAGACAGCAGC 2552

## RESULT 5

US-09-947-199-9  
; Sequence 9, Application US/09947199  
; Patent No. US20020127684A1  
; GENERAL INFORMATION:  
; APPLICANT: Raju, Jeyaseelan  
; TITLE OF INVENTION: NOVEL DARK PROTEIN AND NUCLEIC ACID MOLECULES AND USES  
; TITLE OF INVENTION: THEREFOR  
; FILE REFERENCE: MNI-068CP2  
; CURRENT APPLICATION NUMBER: US/09/947,199  
; PRIOR FILING DATE: 2001-09-05  
; PRIOR APPLICATION NUMBER: 60/111,938  
; PRIOR FILING DATE: 1998-12-11  
; PRIOR APPLICATION NUMBER: 09/291,839  
; PRIOR FILING DATE: 1999-04-14  
; PRIOR APPLICATION NUMBER: 09/458,457  
; PRIOR FILING DATE: 1999-12-10  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 9  
; LENGTH: 2505  
; TYPE: DNA  
; ORGANISM: Rattus norvegicus  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(2505)  
US-09-947-199-9

Query Match 75.5%; Score 1892.2; DB 9; Length 2505;  
Best Local Similarity 84.7%; Pred. No. 0;  
Matches 2122; Conservative 0; Mismatches 383; Indels 0; Gaps 0;

Qy 1 ATGGGAAATTTACAAATCCAGACCAAACACAGACTTTGTTGATGAATGGAAAGAAAGTT 60

Db 1 ATGGGAAATTTACAAATCCAGACCAAACACAGACTTTGTTGATGAATGGAAAGAAAGTT 60  
Qy 61 AGTGAATCATATGTTATPCAAATAGAAAGATTAGAAGATGACCTGCAGATCAAGGAAAAA 120  
Db 61 AGTGAATCTTACCTATATCATAGAAAGCTGGAGATAAAGCTGCAGATCAAGGAAAT 120  
Qy 121 GAACAGACAACTAAGGAATATATTTGGCTCTGATGAAGCCCTTCAAGTAAAGTCAATTTA 180  
Db 121 GAATTTCAAGAACTAAGGCACATCTTTGGCTCTGATGAAGCCCTTCAAGTAAAGTCAATTTA 180  
Qy 181 AATTACCGCACTGAAATGGCTGTCTACTTCTATTTATTTGTTGCAATTTGTTGGAGCAAG 240  
Db 181 AATTACCGCACTGAAATGGCTGTCTACTTCTATTTATTTGTTGCAATTTGTTGGAGCAAG 240  
Qy 241 AATTCACATATTCGAATCTTATGTTGAAAGGGCTCCGCCATTCCTGCATGACAAAGAAAT 300  
Db 241 AAGTCACATATTCGCTGCTTATGTTTAAAGGGCTCCGCCATTCCTGCATGACAAAGAAAT 300  
Qy 301 GGATTTACAGCTTGCAATTTAGCAGTTTACAGGAATAATGCAGAAATTTGATCACTTCTCTG 360  
Db 301 GGGTTTCCAGCTCTGACCTGGCCGTTTACAGGACAGCCCGGAATTTATCACTTCTACCTG 360  
Qy 361 CTTACAGTGGAGCTGATATACAGAGGTTGGATACGGTGGCTCACTGCCCTCCATATTT 420  
Db 361 TTGCACAGCGAGCAGATGTTACGAAGTGGGATACGGTGGCTCACTGCCCTCCACATA 420  
Qy 421 GCTACATAGCTGGCCACCTAGAGGCTGCTGATGCTGTTTGCACATGAGCTATATGTC 480  
Db 421 GCTGCAATAGCTGGACACCCAGAGGCTGCAGAGTGTCTCTACAACATGGGCCAATGTC 480  
Qy 481 AATATTCAGATGCGCTTTTTCACCTCCATTCGATTCGATTCGAGCGTACTATGGAATGAA 540  
Db 481 AATGTTCAAGATGCGCTTCTTTCACCCCATTCGACATTCGAGCGTACTATGCGCACAG 540  
Qy 541 CAGTAACTCGCTTCTTTTGAATTTGGTGTGATGTAATGTAAGTGGTGAAGTTGA 600  
Db 541 CAGTAACTCGCTTCTTTTGAATTTGGTGTGATGTAATGTAAGTGGTGAAGTTGG 600  
Qy 601 GATAGACCCCTCCACCTAGCATCTGCAAAAGGATTCCTTGAATATTCGCAAACTCTTGATG 660  
Db 601 GACAGGCTCTGCACCTGGCTCTGCAAGGGCTTCTTCAACATTTGAAACTCTTGTA 660  
Qy 661 GAAGAAGGACGCAAAAGCAGATGTGAATGCTCAAGATAATCAAGACCATTCCTCCATCCAT 720  
Db 661 GAAGAAGGACGCAAAAGCAGATGTGAACGCTCAGGACAAATGAAGACCCACCTCTCTGCAC 720  
Qy 721 TTCTGTTCTCGATTTGGACACCATGATATAGTAAATATCTGCTGCAAGTGAATTTGAA 780  
Db 721 TTCTGTTCTCGATTTGGACACCAATATAGTGAAGTACTCTCTCCAGAGTGACTTAGAG 780  
Qy 781 GTTCAACCTCATGTTGTTAATATCTATGGAGATACCCCTTACACCTGGCATGCTACAT 840  
Db 781 GTCCAGCTCAAGTCATTAACATCTATGGTGACATCTCTTTGCACCTGGCATGCTACAT 840  
Qy 841 GGCAAAATTTGAAGTGGCAAGGAAATCATCCAAATATCAGGAACAGAAAGTCTGACTAAG 900  
Db 841 GGAAATTTTGAAGTGGCAAGGAAATTTGTCAGGTAAACAGGAATCTGAAAGTCTGACTAAG 900  
Qy 901 GAAACACATCTTCAAGTGAACACAGCTTTTCAAGTGTGTGCTTATGCTGCAAGAGCTTAC 960  
Db 901 GAAACACATCTTCAAGTGAACACAGCTTTTCAAGTGTGTGCTTATGCTGCAAGAGCTTAC 960  
Qy 961 CTAGTCAAAATTTCTTCTGATCAGAAATGTATGAATCAACCAACCAAGGAGGATGG 1020  
Db 961 CTGTCAAAATTTCTTCTGATCAGAAATGTGTGAACATTAACCAACCGAGGAGAGATGG 1020  
Qy 1021 CACACTGGATTAACATCTGCTTGTCTACCAAGGTCACATTCGGCTGGTTCAGTCTTACTG 1080  
Db 1021 CACACAGGATTCAGCTCTGCTTGTCTACCAAGGTCATTCGGCTGGTTCAGTCTTACTT 1080  
Qy 1081 GATAATGGAGCTGATATGAATCTAGTGGCTTGTGATCCCGAGGAGGCTAGTGGTGAATAA 1140  
Db 1081 GATAATGGTGCAGATATGAATCTTGTGCTTGTGATCCCGAGGAGGCTAGTGGTGAATAA 1140

```
QY 1141 GATGAGCAGCATGTTTGTGATGTGGCTTATGAATAAGGCGATGATGCCATTGTCACTC 1200
Db 1141 GATGAGCAGCATGTTTGTGATGTGGCTTACGAGAAGGACATGATGCAATTTGTACACTC 1200
QY 1201 CTGAAGCATTAAGAGACCAACAGATGAATGGCCCTGTATGTAATGAATATCTCAGCCTGGA 1260
Db 1201 CTGAAGCATTAAGAGACCAACAGAGAGCTGCCATGTAACGAATATTTCCAGCCTGGA 1260
QY 1261 GGAGATGGCTCCTATGTGTCTGTCTCCATCACCTTGGGGAAGATTAAGAAAGCATGACAAAA 1320
Db 1261 GGAGATGGCTCCTATGTGTCTGTCTCCATCACCTTGGGGAAGATTAAGAAAGCATGACAAAA 1320
QY 1321 GAGAAGGAGATATTTCTCTCTCTAAGAGCTGGATGGCTTCCATTTCCATTTCCATTCAGCTC 1380
Db 1321 GAGAAGGAGATATTTCTCTCTCTAAGAGCTGGATGGCTTCCATTTCCATTTCCATTCAGCTC 1380
QY 1381 TCAGAAATGAGTTCATGAGATTAATGGCTCAGGTTCTTTGGGAAAGATATATAAGGA 1440
Db 1381 TCCGAAATCGAGTTCACGAGATTAATGGCTCGGGTTCCTTTGGGAAAGTCTATAAGGG 1440
QY 1441 CGATGAGAAATAAATAGTGGCTATAAAGCTTATGAGCCAAATACCTACTGCTCCAAG 1500
Db 1441 CGATGAGAAATAAATAGTGGCTATAAAGCTTATGAGCCAAATACCTACTGCTCCAAG 1500
QY 1501 TCAGATGTGGATATGTTTGGCGAGAGGTGCCATTTCTGCCAGCTCAATCATCCTCTGC 1560
Db 1501 TCAGATGTGGATATGTTTGGCGAGAGGTGCCATTTCTGCCAGCTCAATCATCCTCTGC 1560
QY 1561 GTAATTCAGTTTGGGTGCTTGTGTTGAATGATCCAGCCAGTTTGGCATTTGCACTCAA 1620
Db 1561 GTGGTTTCAGTTTGGGTGCTTGTGTTGAATGATCCAGCCAGTTTGGCATTTGCACTCAA 1620
QY 1621 TACATATCAGGGGTTCTCTGTTCTCCCTCTCTCATGAGCAGAGAGATTTCTGATTTG 1680
Db 1621 TACATATCAGGGGTTCTCTGTTCTCCCTCTCTCATGAGCAGAGATTTCTGATTTG 1680
QY 1681 CAGTCTAAATTAATTTGCAAGTATGATTTGCCAAGGCATGGAGTACCTTCAACCTTG 1740
Db 1681 CAGTCTAAATTAATTTGCAAGTATGATTTGCCAAGGCATGGAGTACCTTCAACCTTG 1740
QY 1741 ACAGAGCAATATACATGCTGATTTGAGAGTACATATTTCTCTATGAGGATGG 1800
Db 1741 ACCAGCAATATACATGCTGATTTGAGAGTACATATTTCTCTATGAGGATGG 1800
QY 1801 CATGCTGTGGTGGCAGATTTGGAGATCAAGATTTCTACAGTCTCTGGATGAAGACAAC 1860
Db 1801 CATGCTGTGGTGGCAGATTTGGAGATCAAGATTTCTACAGTCTCTGGATGAAGACAAC 1860
QY 1861 ATGACAAACAACTGGGAACTCGTGTGGATGGCTCTGAGGTGTTTCAAGAGTGCAT 1920
Db 1861 ATGACAAACAACTGGGAACTCGTGTGGATGGCTCTGAGGTGTTTCAAGAGTGCAT 1920
QY 1921 CGGTACACATCAAGCAGATGTTTCAAGTATGCTGTCTGTCTGTGGGAAATTTCTCACT 1980
Db 1921 AGATACACATCAAGCAGATGTTTCAAGTATGCTGTCTGTCTGTGGGAGTCTCTCACT 1980
QY 1981 GCGGAAATTCATTCGCTCATCTCAAGCCAGCGCTCGGAGCAGACATGGCTTACCAC 2040
Db 1981 GGAGAAATTCATTCGCTCATCTCAAGCCAGCGCTCGGAGCAGACATGGCTTACCAC 2040
QY 2041 CACATCAGACCTCCATTTGGTATTTCCATTTCCAGGCCATATCATCTCTGCTGATACGA 2100
Db 2041 CACATCAGACCTCCATTTGGTATTTCCATTTCCAGGCCATATCATCTCTGCTGATACGA 2100
QY 2101 GGGTGGAGCGCATGTTCTCAAGGAGAGCCGAAATTTTCTGAAGTTGTCAAGAGTTAGAA 2160
Db 2101 GGGTGGAGCGCATGTTCTCAAGGAGAGCCGAAATTTTCTGAAGTTGTCAAGAGTTAGAA 2160
QY 2161 GAGTGTCTCTGCAACATTTGAGCTGATGTTCTCTGATCAAGTACAGAGTGGTCTCTC 2220
Db 2161 GAGTGTCTCTGCAACATTTGAGCTGATGTTCTCTGATCAAGTACAGAGTGGTCTCTC 2220
```

## RESULT 6

```
US-10-626-173-9
; Sequence 9, Application US/10626173
; Publication No. US20040110232A1
; GENERAL INFORMATION:
; APPLICANT: Raju, Jeyaseelan
; TITLE OF INVENTION: NOVEL CARP PROTEIN AND NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: MMI-068CP2
; CURRENT APPLICATION NUMBER: US/10/626,173
; PRIOR FILING DATE: 2003-07-24
; PRIOR APPLICATION NUMBER: US/09/947,199A
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: US 09/291,839
; PRIOR FILING DATE: 1999-04-14
; PRIOR APPLICATION NUMBER: US 09/458,457
; PRIOR FILING DATE: 1999-12-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 9
; LENGTH: 2505
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2505)
US-10-626-173-9
```

```
Query Match 75.5%; Score 1892.2; DB 17; Length 2505;
Best Local Similarity 84.7%; Pred. No. 0;
Matches 2122; Conservative 0; Mismatches 383; Indels 0; Gaps 0;
```

```
QY 1 ATGGGAATTAATACTAGACCAACCCAAACCTTTACTGATGAATGGAGAAAAGTC 60
Db 1 ATGGGAATTAATACTAGACCAACCCAAACCTTTACTGATGAATGGAGAAAAGTC 60
QY 61 AGTGAATCATATGTTATCAACATAGAAAGATTAGAAGATGACCTGCAGATCAAGGAAA 120
Db 61 AGTGAATCATATGTTATCAACATAGAAAGCTGGAGATAAAGCTGCAGATCAAGGAAA 120
QY 121 GAATGACAGCACTTAGGATATATTTGGCTCTGATGAAGCCCTTCAGTAAAGTCAATT 180
Db 121 GAATGACAGCACTTAGGATATATTTGGCTCTGATGAAGCCCTTCAGTAAAGTCAATT 180
QY 181 AATTACCGCACTGAAAATGGGCTGCTCTACTTTTATTTTGTTCATTTTGGAGGCAAG 240
Db 181 AATTACCGCACTGAAAATGGGCTGCTCTACTTTTATTTTGTTCATTTTGGAGGCAAG 240
QY 241 AATACATATTCGAATCTTATGTTGAAGGGCTCGGCCATCTCGACTGACAGAAAT 300
Db 241 AATACATATTCGAATCTTATGTTGAAGGGCTCGGCCATCTCGACTGACAGAAAT 300
```

QY 301 GGATTTACAGCTGCAATTTAGCAGTTTACAAGGATAATGCAGAAATGATGATCACTTCTCTG 360  
Db 301 GGGTTTCAGCTGCGACCTGGCGGTTTTACAAGGACAGCCGGAATATATCACTTCACTG 360  
QY 361 CTTACAGTGGAGCTGATATACAGCAGGTTGGATAGCGTGGCTCACTGCCCTCCATATT 420  
Db 361 TTGCACAGCGGACAGATGTTACGAACTGGGATACGGTGGCTCACAGCCCTCCACATA 420  
QY 421 GCTACAATAGTGGCCACCTAGAGGCTGCTGATGCTGTGGCAACATGGAGCTAATGTC 480  
Db 421 GCTCAATAGTGGACACCCAGAGGCTGCAGAGTGTCTGCTCAACATGGGCCCAGTGTG 480  
QY 481 AATATTCAAGATGAGTGTGTTTTTCACTCCATTCATATTGCAGCGPACTATGGACATGAA 540  
Db 481 AATGTTCAAGATGCGCTTCTTCAACCCACTGCACATTTGCAGCCTACTATGGGACAGG 540  
QY 541 CAGTAATCTGCCCTTCTTTGAAATTTGGTCTGATGTAATGTAAGTGGTGAAGTTGA 600  
Db 541 CAGTAACCAAGTGTCTTTTGAAGTTTGGTCTGATGTCAATGTAAGCGGTGAAAGTTGG 600  
QY 601 GATAGACCTCCACCTAGCATCTGCAAGAGGATCTTGAATATTGCAAACTCTTGATG 660  
Db 601 GACAGGCTCTGCACCTGGGCTCTGCAAGAGGCTTCTTCAACATTTGTAAGTCTCTGGTA 660  
QY 661 GAAGAAGGACCAAGCAGATGTGAACGCTCAGGACAAATGAAGACCACTGCTCTGCAC 720  
Db 661 TTCTGTTCTGATTTGACACCAATGATAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 780  
QY 721 TTCTGTTCTGATTTGACACCAATGATAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 780  
Db 721 TTCTGTTCTGATTTGACACCAATGATAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 780  
QY 781 GTTCAACCTCATGTTGTAATATCTATGAGATACCCCTTACACCTGCATGCTACAAAT 840  
Db 781 GTCCAGGCTCAGTCAATATCACTATGTTGACACTTCTTGCACCTGCGATGCTACAAAT 840  
QY 841 GGCAAAATTTGAAGTGGCCAAAGGAAATCAATCCAAATATCAGGAACAGAAAGTCTGACTAG 900  
Db 841 GGAAATTTTGAAGTGGCCAAAGGAAATGTTCCAGGTAACAGGAATGAAAGTCTGACTAG 900  
QY 901 GAAACATCTTCAGTGAACAGCTTTTATAGTGTGTTGATGCTGATGCTGCTGCTGCTGCTG 960  
Db 901 GAAACATCTTCAGTGAACAGCTTTTATAGTGTGTTGATGCTGATGCTGCTGCTGCTGCTG 960  
QY 961 CTAGTCAAAATTTCTGTTGATCAGAAATGCTAATAACATCAACCAAGGAGGAGTGGG 1020  
Db 961 CTGGTCAAAATTTCTGTTGATCAGAAATGCTGTAACATTAACCAAGGAGGAGTGGG 1020  
QY 1021 CACACTGATTAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080  
Db 1021 CACACAGGATTTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080  
QY 1081 GATAATGAGCTGATATGAATCTAGTGGCTGTGATCCAGCAGGCTCTAGTGGTGAATAA 1140  
Db 1081 GATAATGAGCTGATATGAATCTAGTGGCTGTGATCCAGCAGGCTCTAGTGGTGAATAA 1140  
QY 1141 GATGAGCAGACATGTTTGTGTTGGCTTTATGAAAGGAGGAGGATGATGCCATTTGTACACTC 1200  
Db 1141 GATGAGCAGACATGTTTGTGTTGGCTTTATGAAAGGAGGAGGATGATGCCATTTGTACACTC 1200  
QY 1201 CTGAAGCATATTAAGACACCAAGATGAAATTCGCCCTGTAATGAATATTTCTCAGCTGGA 1260  
Db 1201 CTGAAGCATATTAAGACACCAAGATGAAATTCGCCCTGTAATGAATATTTCTCAGCTGGA 1260  
QY 1261 GGAGATGGCTCCTATGTTGTTGTTTCCATCAACCTTGGGGAAGATTTAAAGCATGACAAA 1320  
Db 1261 GGAGATGGCTCCTATGTTGTTTCCATCAACCTTGGGGAAGATTTAAAGCATGACAAA 1320  
QY 1321 GAGAGGAGATATTTCTCTCTTCAAGAGCTGGATGCTTCAATTTCCATCTTCTGCTC 1380  
Db 1321 GAGAGGAGATATTTCTCTCTTCAAGAGCTGGATGCTTCAATTTCCATCTTCTGCTC 1380

QY 1381 TCAGAAATTCAGTTCATGAGATTATTGGCTCAGGTTCTTTTGGGAAAGTATATAAAGA 1440  
Db 1381 TCGAAATTCAGTTCATGAGATTATTGGCTCAGGTTCTTTTGGGAAAGTATATAAAGA 1440  
QY 1441 CGATGCAGAAATTAATAGTGGCTATPAAACGTTATCGAGCCAAATACCTACTGCTCCAAG 1500  
Db 1441 CGATGCAGAAATTAATAGTGGCTATPAAACGTTATCGAGCCAAACACCTACTGCTCCAAG 1500  
QY 1501 TCAGATGTGATATGTTTTCGAGAGGTTGTCATTTCTGCGAGCTCAATCATCCCTGC 1560  
Db 1501 TCAGAGCTGATATGTTTTCGCGAGAGGTTGTCATTTCTGCGAGCTCAACCAACCTCTGC 1560  
QY 1561 GTAAATTCAGTTCCTGGGTTGCTTGAATGATCCAGCCAGTTCCTTGCCTACTCAA 1620  
Db 1561 GTGGTTTCAGTTCCTGGGTTGCTTGAATGATCCAGCCAGTTCCTTGCCTACTCAA 1620  
QY 1621 TACATATCAGGAGGTTCTCTGTTCTCCCTCTCATGACAGAGAGGATTTCTGATTTG 1680  
Db 1621 TACATTTTCAGAGAGGCTCCCTGTTCTCCCTGCTCATGAAACAGAGAGATTTCTGACTG 1680  
QY 1681 CAGTCTAAATTAATATTGTCAGTAGATGTTGCAAGGATGAGTACCTTCAACACCTG 1740  
Db 1681 CAGTCTAAATTAATATTGCGGTAGAGCTTGCAGAGGATGAGTACCTTCAACACCTG 1740  
QY 1741 ACAGAGCAATATATCATGTCACITTTGAGAGTCAAGTCAATATTTCTCTATGAGATGG 1800  
Db 1741 ACCAGGCAATATATCATGTCACITTTGAGAGTCAAGTCAATATTTCTCTATGAGATGG 1800  
QY 1801 CATGCTGTGGTGGCAGATTTTGGAGAAATCAAGATTTCTACAGTCTCTGATGAAGACAA 1860  
Db 1801 CATGCTGTGGTGGCAGATTTTGGAGAAATCAAGATTTCTGAGTCCCTGGATGAAGACAA 1860  
QY 1861 ATGACAAAACCACTGGGAACCTTCGTTGGATGGTCTCTGAGGTGTTACGAGTGCACT 1920  
Db 1861 ATGACAAAACCACTGGGAACCTTCGCTGGATGGTCTCTGAGGTGTTACGAGTGCACT 1920  
QY 1921 CGGTACACCATCAAGAGAGATGTTTACAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1980  
Db 1921 AGATACACCATCAAGAGAGATGTTTACAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1980  
QY 1981 GGGCAAAATTCATTCGCTCATCTCAAGCCAGCGGCTGCGCCAGCAGACATATGCTTACCAC 2040  
Db 1981 GGAGAAATTCATTCGCTCATCTCAAGCCAGCGGCTGCGCCAGCAGACATATGCTTACCAC 2040  
QY 2041 CACATCAGACTCCCATTTGGCTATTCATTTCCCAAGCCCATCTCATCCCTGCTGATCGA 2100  
Db 2041 CACATCAGACTCCCATTTGGCTATTCATTTCCCAAGCCCATCTCATCCCTGCTGATCGA 2100  
QY 2101 GGGTGAACCATGCTCTGAAAGAACCCGAAATTTCTGAAAGTTGCTCATGAAGTTAGAA 2160  
Db 2101 GGGTGAACCATGCTCTGAAAGAACCCGAAATTTCTGAAAGTTGCTCATGAAGTTAGAA 2160  
QY 2161 GAGTGTCTGCAACATTCAGTGTGATGTTCTCTGCTCATCAAGTAAACAGAGTGGGTTCTCT 2220  
Db 2161 GAGTGTCTGCAACATTCAGTGTGATGTTCTCTGCTCATCAAGTAAACAGAGTGGGTTCTCT 2220  
QY 2221 TCACCTTCTTCTTCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2280  
Db 2221 TCACCTTCTTCTTCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2280  
QY 2281 GCAGCATTAAGAAAGTGGTTCGAAATTCGAAATTCGCTTAAATGCAAGGTCCTATGCTGCT 2340  
Db 2281 GCAGCATTAAGAAAGTGGTTCGAAATTCGAAATTCGCTTAAATGCAAGGTCCTATGCTGCT 2340  
QY 2341 TTGTCCCAAGAGTGGTGAACATTTCTCTCAAGGCTGCTCTTTGGAGGAGATGAAGA 2400  
Db 2341 TGSTCCCAAGAGTGGTGAACATTTCTAAATCCGGGCTGCTTTTGGAGGAGATGAATAGG 2400  
QY 2401 AGTCTTCAATACACACCTTGAACATATGCTATGCTATCGATCCCATGAGCTCAATG 2460  
Db 2401 AGCACCAGATTTCAACTGTTGACAAATACGGCTATGCTCTGATCCCATGAGCTGAGC 2460  
QY 2461 CATTTTCATTTCTTCCCAAAATAGTAGCAGCTTTGAGGACAGCAGC 2505



Db 1681 TACATTTTCAGGAGGCTCCCTGTTCTCCCTGCTTCATGAACAGAGAGAAATCTTGACTTG 1740  
Qy 1681 CAGCTCAAAATTAATTTTCAGTAGATGTTGCCAAGGCGATGAGTACTCTTACAACTG 1740  
Db 1741 CAGCTCAAAATTAATTTTCAGTAGATGTTGCCAAGGCGATGAGTACTCTTACAACTG 1800  
Qy 1741 ACACAGCAATTAATACATGCTGATGTTGAACAGTCAAAATTTCTTCTCTATGAGGATGG 1800  
Db 1801 ACCAGCAATATACATGCTGATGTTGAACAGTCAAAATTTCTTCTCTATGAGGATGG 1860  
Qy 1801 CATGCTGTGTGAGATTTTGGAGATCAAGATTTCTACAGTCTCTGATGAGTCAAGCAAC 1860  
Db 1861 CATGCTGTGTGAGATTTTGGAGATCAAGATTTCTACAGTCTCTGATGAGTCAAGCAAC 1920  
Qy 1861 ATGACAAACAACTGGGAACCTCCGTTGGATGGCTCCCTGAGTGTTCACGAGTGCAC 1920  
Db 1921 ATGACAAAGCAGCAGGAACCTCCGTTGGATGGCTCCCTGAGTGTTCACAGTGCAC 1980  
Qy 1921 CGGTACACCATCAAGCAGATGCTTTCAGTATGCTCTGTGTCTGTGGGAATTTCTCACT 1980  
Db 1981 AGATACACCATCAAGCAGTATGCTTTCAGTATGCTCTGTGTCTGTGGAGTCTCTCACT 2040  
Qy 1981 GGGGAATTCATTCGCTCATCTCAAGCAGGCTGCGGAGCAGACATGCTTACCAAC 2040  
Db 2041 GGAAGAAATTCATTCGCTCATCTCAAGCAGGCTGCGGAGCAGATGAGCGTATCAC 2100  
Qy 2041 CACATCAGACCTCCCATTTGGCTATTCATTCACAGCCCATATCATCTCTGTGATCGA 2100  
Db 2101 CACATCAGACCGCCATTCGCTATTCATTCACAGCCCATATCATCTCTGTGATCGG 2160  
Qy 2101 GGGTGAACGATGCTCTGAGAGAGACCGCAATTTCTGAGTGTCTCATGAGTAGAA 2160  
Db 2161 GGGTGAATGCATGCTCTGAGAGACCGCAATTTCTGAGTGTCTCATGAGTAGAA 2220  
Qy 2161 GAGTGTCTCTGCAACATGAGCTGATGCTCTCTGATCAAGTCAAGTCAAGTCAAGTCAAG 2220  
Db 2221 GAGTGTCTCTGCAACATGAGCTGATGCTCTCTGATCAAGTCAAGTCAAGTCAAGTCAAG 2280  
Qy 2221 TCACCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2280  
Db 2281 TCACCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2340  
Qy 2281 GCACATTAAGAGTGTCTCTGAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2340  
Db 2341 GCACATTAAGAGTGTCTCTGAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2400  
Qy 2341 TTGTCCAAAGTGTGACAAATATTCCTCTCAAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCT 2400  
Db 2401 TGGTCCAAAGTGTGACAAATATTCCTCTCAAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCT 2460  
Qy 2401 AGCTTCAATACACACCATGACAAATATTCCTCTCAAGTGTCTCTCTCTCTCTCTCTCTCTCT 2460  
Db 2461 AGCACCAGTATCACTGTGTGACAAATATTCCTCTCAAGTGTCTCTCTCTCTCTCTCTCTCTCT 2520  
Qy 2461 CATTTTCATTTCTTGGCAATATGATGAGCTTTTGGAGCAGCAGC 2505  
Db 2521 CACCTTCACTCCCGCAAGCAGCAGCACTTTTGGAGCAGCAAC 2565

## RESULT 8

US-10-626-173-7  
; Sequence 7, Application US/10626173  
; Publication No. US20040110232A1  
; GENERAL INFORMATION:  
; APPLICANT: Raju, Jeyaseelan  
; TITLE OF INVENTION: NOVEL DARK PROTEIN AND NUCLEIC ACID MOLECULES AND USES  
; FILE OF INVENTION: THEREFOR  
; FILE REFERENCE: MNI-068CP2  
; CURRENT APPLICATION NUMBER: US/10/626,173  
; PRIORITY FILING DATE: 2003-07-24  
; PRIOR APPLICATION NUMBER: US/09/947,199A  
; PRIOR APPLICATION NUMBER: US 60/111,938  
; PRIOR FILING DATE: 1998-12-11

; PRIOR APPLICATION NUMBER: US 09/291,839  
; PRIOR FILING DATE: 1999-04-14  
; PRIOR APPLICATION NUMBER: US 09/458,457  
; PRIOR FILING DATE: 1999-12-10  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 7  
; LENGTH: 3026  
; TYPE: DNA  
; ORGANISM: Rattus norvegicus  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (61)..(2565)  
US-10-626-173-7

Query Match 75.5%; Score 1892.2; DB 17; Length 3026;  
Best Local Similarity 84.7%; Pred. No. 0;  
Matches 2122; Conservative 0; Mismatches 383; Indels 0; Gaps 0;  
Qy 1 ATGGGAAATTAATAATCTAGACCAACCCAACTTGTACTGATGAATGGAAGAAAAGTC 60  
Db 61 ATGGGAAATTAATAATCTAGACCAACCCAACTTGTACTGATGAATGGAAGAAAAGTT 120  
Qy 61 AGTGAATCATATGTTATCAATAGAAAGATTAGAGATGACCTCAGATCAAGGAAAA 120  
Db 121 AGTGAATCTTACGCTATATCATGAAAGGCTGGAGATACTGCAGATCAAGGAAAT 180  
Qy 121 GAATCAACAACTAAGGAATATATTTGGCTCTGATGAAGCTTCAGTAAAGTCAATTTA 180  
Db 181 GAATTTCAAGAACTAAGGCACATCTTTGGCTCTGATGAAGCTTCAGTGAAGTCAATTTA 240  
Qy 181 AATACCGCACTGAAATGGGCTGCTCTACTCTTATTTATGTTGATTTGTTGGAGGCAAG 240  
Db 241 AATACCGCAGACAGGCTGGGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 300  
Qy 241 AATATCATATTTGCAACTCTTATGTTGAAAGGGCTCCGCCATCTCGACTGCAAGAAAT 300  
Db 301 AAGTCACATATCGTGCCCTTATGTTTAAAGGGCTCCGTCCTCAGACTGACGAGAAAT 360  
Qy 301 GGAATTCAGACCTTGATTTAGCAGTTTACAGGATTAAGAGATTAAGCAGATTTGATCTCTG 360  
Db 361 GGGTTTTCCAGCTCTGACCTGCGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 420  
Qy 361 CTTCACAGTGGAGCTCATATACAGAGGTTGATACGGTGGCTCACTGCGCTCCATATTT 420  
Db 421 TTGCACAGCGAGCAGATGTTGAGGAGTGGATACGGTGGCTCACAGCCCTCCACATA 480  
Qy 421 GCTACAAATAGCTGGCCACCTAGAGGCTGCTGATGCTGCTTTGCAACATGAGAGCTAATGTC 480  
Db 481 GCTGCAATAGCTGGACACACCCAGAGGCTGCAGAGTGTCTCTACAACATGGGGCCAAATGTG 540  
Qy 481 AATATTCAGATGCACTTTTCTACTCCATTCATATTCAGCTGAGCTGCTATGACATGAA 540  
Db 541 AATGTTCAAGATGCGCTCTCTTCCACCCAGCTGCAATTCAGAGCTCTATGCGCAGAG 600  
Qy 541 CAGGTAACCTGCTCTTTTGAATTTGGTGTGATGTAATGTAAGTGGTGAAGTTGGA 600  
Db 601 CAGGTAACCTGCTCTTTTGAATTTGGTGTGATGTAATGTAAGTGGTGAAGTTGGA 660  
Qy 601 GATAGACCCCTCCACCTAGATCTGCAAAAGGATCTTGAATATTCGAAACTCTTGATG 660  
Db 661 GACAGGCTCTGACCTGGCTCTGCAAAAGGCTTCTTCAACATTTGTGAATCTCTGGTA 720  
Qy 661 GAAGAGGCGACAAAGCAGATGTAATGCTCAAGATTAATGAAGACATGTCCTCCATCCAT 720  
Db 721 GAAGAGGCGACAAAGCAGATGTAATGCTCAAGATTAATGAAGACATGTCCTCTGAC 780  
Qy 721 TTCTGTTCTCGATTTGGACACCAATGATATAGTAAAGTCTCTCTGCAAGTGAATTTGAA 780  
Db 781 TTCTGTTCTCGATTTGGACACCAATGATATAGTAAAGTCTCTCTGCAAGTGAATTTGAA 840  
Qy 781 GTTCAACCTCATGTTGTTAATATCTATGAGATACCCCTTACCTGCGATGCTACAT 840

Db 841 GTCCAGCCTCAGCTCAATTAACATCTATGGTGACACTCCTTTGGACCTGGCATGCTACAAT 900  
Qy 841 GCGAAATTTGAAGTTGCCAAGGAATCATCCAAATATCAGAAACAGAAAGTCTGACTAAG 900  
Db 901 GCGAAATTTGAAGTTGCCAAGGAATTTGCCAGGTAAACAGGAACCTGAAGTCTGACTAAG 960  
Qy 901 GAAACATCTTCAGTGAACAGCTTTTCATAGTCTGCTGTACCTATGCGAAGACATTTGAC 960  
Db 961 GAAACATCTTCAGGCGACAGCTTTTCATAGTCTGCTGTACCTATGCGAAGACATTTGAC 1020  
Qy 961 CTAGTCAAAATTTCTTTGATCAGAAATGCTATAAACATCAACCAAGGAAGGATGGG 1020  
Db 1021 CTGGTCAAAATTTCTTTGATCAGAAATGCTGTGAACATTAACACCGGAGAGATGGG 1080  
Qy 1021 CACATGGATTAACATCTGCTGTACCAAGGTACATTTGCCCTGGTTCAGTTCTTACTG 1080  
Db 1081 CACACAGGATTCACATCTGCTGTACCAAGGTACATTTGCCCTGGTTCAGTTCTTACTG 1140  
Qy 1081 GATAATGGAGCTGATATCAATCTAGTGGCTTTGATCCCAAGGCTAGTGGTGAAAAA 1140  
Db 1141 GATAATGGTGACAGATATGAATCTTGTGCTGTGATCCAGAGGCTAGTGGTGAAAAA 1200  
Qy 1141 GATGACAGACATGTTGATGTGGCTTATGAAAAGGCGATGATGCCATTTGCACATC 1200  
Db 1201 GATGACAGACATGTTGATGTGGCTTATGAAAAGGCGATGATGCCATTTGCACATC 1260  
Qy 1201 CTGAAGCATTTAAGAGACCAAGATGAATTCGCTGTAATGAATTTCTCAGCCTGGA 1260  
Db 1261 CTGAAGCATTTAAGAGACCAAGGAGGCTGATGAACGATTTCCAGGCTGGA 1320  
Qy 1261 GGAGATGGCTCCTATGCTGCTTCCATACCCCTGGGGAAGATTAAAGCATGACAAAA 1320  
Db 1321 GGAGATGGCTCCTATGCTGCTTCCATACCCCTGGGGAAGATTAAAGCATGACAAAA 1380  
Qy 1321 GAGAGGCGAGATATCTCCTCTAAGAGCTGATGCTTCCATTTCCATCTTCAGCTC 1380  
Db 1381 GAGAGGCGAGATATCTCCTCTAAGGCTGATGCTTCCATTTCCATCTTCAGCTC 1440  
Qy 1381 TCAGAAATTTGAGTTCCATGAGATTTAGGCTCAGGTTCTTTGGGAAGTATATAAGGA 1440  
Db 1441 TCCGAAATCGAGTTCCACAGAGATTTATCGGCTCGGTTCCCTTTGGGAAGTCTATAAGGG 1500  
Qy 1441 CGATCCGAAATTAATAGTGGCTATATAAGCTTATCGAGCAATACCTACTCTCTCAAG 1500  
Db 1501 CGATCCGAAATTAATAGTGGCTATATAAGCTTATCGAGCAATACCTACTCTCTCAAG 1560  
Qy 1501 TCAGATGTGGATATGTTTTCGAGAGGTTCCATCTCTGCCAGCTCAATCATCCTCTGC 1560  
Db 1561 TCAGAGCTGGATATGTTTTCGAGAGGTTCCATCTCTGCCAGCTCAATCATCCTCTGC 1620  
Qy 1561 GTAATTCAGTTTGGGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1620  
Db 1621 GTGGTTTCAGTTTGGGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1680  
Qy 1621 TACATATCAGGGGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1680  
Db 1681 TACATTTAGAGGCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1740  
Qy 1681 CAGTCTAAATTAATTTAGTATGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1740  
Db 1741 CAGTCTAAATTAATTTAGTATGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800  
Qy 1741 ACACAGCAATTTATACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800  
Db 1801 ACCAGCAATTTATACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1860  
Qy 1801 CATGCTGTGGTGGCAGATTTTGGAGAATCAAGATTTCTACAGTCTCTGATGAAGACAAC 1860  
Db 1861 CATGCTGTGGTGGCAGATTTTGGAGAATCAAGATTTCTGAGTCTCTGATGAAGACAAC 1920  
Qy 1861 ATGACAAAACCTGGGAACCTCGTTGGATGGCTCTGAGTGTTCACGAGTGCATC 1920  
Db 1921 ATGACAAAACCTGGGAACCTCGTTGGATGGCTCTGAGTGTTCACAGTGCACG 1980

Qy 1921 CGGTACACCATCAAGACAGATGCTTTGAGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1980  
Db 1981 AGATACACCATCAAGGCTGATGCTTTGAGTACTCCCTGCTGCTGCTGCTGCTGCTGCT 2040  
Qy 1981 GCGAAATTTCCATTCGCTCATCTCAAGCCAGCGGCTGCGGACAGACATGGCTTACAC 2040  
Db 2041 GCGAAATTTCCATTCGCTCATCTCAAGCCAGCGGCTGCGGACAGATATGGCTATCAC 2100  
Qy 2041 CACATCAGACTCCATTTGCTATTCCATTTCCCAAGCCCATATCATCTCTGCTGATAGA 2100  
Db 2101 CACATCAGAGCGGCCATTCGCTATTCCATTTCCCAAGCCCATCTCATCTCTGCTGATAGG 2160  
Qy 2101 GGGTGGAAACGATGCTGCTGAAGACAGCCGAAATTTCTGAAGTTGCTCATGAAGTTAGAA 2160  
Db 2161 GGGTGGAAATGCTGCTGCTGAAGACAGCCGAAATTTCTGAAGTTGCTCATGAAGTTAGAA 2220  
Qy 2161 GAGTGTCTCTGCAACATTTGAGCTGATGCTCTCTGCAATCAAGTAACAGAGTGGTCTCTC 2220  
Db 2221 GAGTGTCTCTGCAACATTTGAGCTGATGCTCTCTGCAATCAAGTAACAGAGTGGTCTCTC 2280  
Qy 2221 TCACCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2280  
Db 2281 TCACCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2340  
Qy 2281 GCAGCATTAAGAAAGTCTTTTGAATTTGGAATATGCTCTTAAATGCAAGTCTCTATGCTCT 2340  
Db 2341 GCAGCATTTACGAGCGGCTTTTGAATTTGGAATATGCTCTTAAATGCAAGTCTCTATGCTCT 2400  
Qy 2341 TTGTCCCAAGTCTGCAACATTTCTCTCTCAAGTCTCTCTTGGAGAGATGAAAGA 2400  
Db 2401 TGGTCCCAAGTCTGCAACATTTCTCTCTCAAGTCTCTCTTGGAGAGATGAAAGA 2460  
Qy 2401 AGTCTTCAATACACACCATTTGCAAAATATGCTATGCTATGCTATGCTATGCTATGCT 2460  
Db 2461 AGCACCAGTATTTCAACTGTTGACAAATACGCTATGCTCTGATCCCATGAGCTGAG 2520  
Qy 2461 CATTTTCATTTCTGCGCAATATAGTACAGCTTTTGGAGACGAGC 2505  
Db 2521 CACCTTCACTCCGCGCAAGACGACGACCACTTTTGGAGACGAGCAAC 2565

## RESULT 9

US-09-833-381-1149  
; Sequence 1149, Application US/09833381  
; Patent No. US20020132090A1  
; GENERAL INFORMATION:  
; APPLICANT: Robison, Keith E.  
; TITLE OF INVENTION: No. US20020132090A1: Nucleic Acid and Protein Homologs  
; FILE REFERENCE: 5800-119  
; CURRENT APPLICATION NUMBER: US/09/833,381  
; CURRENT FILING DATE: 2001-04-11  
; PRIOR APPLICATION NUMBER: 09/516,448  
; PRIOR FILING DATE: 2000-02-29  
; NUMBER OF SEQ ID NOS: 2050  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1149  
; LENGTH: 616  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-833-381-1149

Query Match 17.1%; Score 427.4; DB 9; Length 616;  
Best Local Similarity 81.8%; Pred. No. 4.3e-119;  
Matches 505; Conservative 0; Mismatches 111; Indels 1; Gaps 1;

Qy 291 GACAGAAATTTGAGTTTACAGCTTGCATTTAGCAGTTTACAGGATTAATGCAAGATTTGAT 350  
Db 1 GACAGAAATTTGAGTTTCCAGCTCTGACCTGCGCGCTTTACAAAGGACAGCCCGAACTTAT 60  
Qy 351 CACTTCTCTCTTTCAGAGTGGAGCTGATATACAGAGGTTGATACGCTGGCTCTCCTCCTGC 410  
Db 61 CACTTCTCTCTTTCAGAGGAGGAGATGTTTTCAGCAAGTGGATACGCTGGCTCTCCTCCTGC 120

Qy 411 CCTCCATATTGCTACATAGCTGGCCACCTAGAGGCTGCTGATGTGCTGTTTGCACATGG 470  
Db 121 CCTCC-CATAGCTGCAATAGCTGGACACCCAGAGGCTGCAGAAGTGCTGCTGCAACATGG 179  
Qy 471 AGCTAATGTCATATTCAAGATGCAAGTATTTTTCACCTCCATTCATATTCGAGCGTACTA 530  
Db 180 GGCCAAAGCTGAATGTTCAAGATGCCGTCTTCTTCACCCACATTCGACGCTACTA 239  
Qy 531 TGGACATGAACAGGTAAGTCCCTCTCTTTTGAATTTGGTGTGATGTAATATGTAAGTGG 590  
Db 240 TGGCAGCAGCAGGTAACCAAGTGCCTTTTGAATTTGGTGTGATGTAATATGTAAGCGG 299  
Qy 591 TGAAGTTGGAGATAGAGCCCTCCACCTAGCATCTGCAAAAGATTTCTGAATATTCGAAA 650  
Db 300 TGAAGTTGGGACAGGCTCTGCACCTGGCCTCTGCAAAAGGCTCTCTTCAACATTTGTAA 359  
Qy 651 ACTCTGATGAAGAGGAGGAGCAAGCAGATGTAATGCTCAAGATAATGAAGACCATGT 710  
Db 360 ACTCTGTAGAGAGGAGAGCAAGCAGATGTAATGCTCAAGATAATGAAGACCATGT 419  
Qy 711 CCACATCCATTTCTGTTCTCGATTTGGACACCATGATATAGTTAAGTATCTCTGCTGAAAG 770  
Db 420 CCCTCTGCACTTCTGTTCTCGATTTGGACACCATGATATAGTTAAGTATCTCTGCTGAAAG 479  
Qy 771 TGAATTTGGAAGTTCACCTCATGTTGTTAATATCTATGGAGATACCCCTTACACCTGGC 830  
Db 480 TGACTTAGAGTCCAGGCTCACGTCATTAACATCTATGTTGACATCTCTTTCGACCTGGC 539  
Qy 831 ATGTACAATGGCAAAATTTGAAGTTGCCAAGGAAATCATCCAATATCAGGAACAGAAAG 890  
Db 540 ATGTACAATGGAAATTTTGAAGTTGCCAAGGAAATTTCCAGGTAACAGGAACGTAAAG 599  
Qy 891 TCTGACTAAGGAAACA 907  
Db 600 TCTGACTAAGGAAACA 616

## RESULT 10

US-09-833-381-1150  
; Sequence 1150, Application US/09833381  
; Patent No. US20020132090A1  
; GENERAL INFORMATION:  
; APPLICANT: Robison, Keith E.  
; TITLE OF INVENTION: No. US20020132090A1e1 Nucleic Acid and Protein Homologs  
; FILE REFERENCE: 5800-119  
; CURRENT APPLICATION NUMBER: US/09/833,381  
; PRIOR FILING DATE: 2001-04-11  
; PRIOR APPLICATION NUMBER: 09/516,448  
; NUMBER OF SEQ ID NOS: 2050  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1150  
; LENGTH: 736  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)...(736)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-833-381-1150

Query Match 16.6%; Score 414.6; DB 9; Length 736;  
Best Local Similarity 81.2%; Pred. No. 4e-115;  
Matches 480; Conservative 0; Mismatches 111; Indels 0; Gaps 0;  
Qy 265 TTGAAGGGCTCCGCCCATCTCGACTGACAGAAATGGATTTACAGCCTTGATTTAGCA 324  
Db 1 TTAAGGGCTCCGCTCCATCCAGACTCAGAGAAATGGGTTTCCAGCTCTGACCTGGCC 60  
Qy 325 GTTTACAGGATAATGCAAGTATGATCACTCTCTCTGCTTCCACAGTGGAGCTGATATACAG 384  
Db 61 GTTTACAGGACACCCCGGAACATATCACTTCACTGTTGTCACAGCGGAGCAGATGTTACAG 120

Qy 385 CAGGTTGGATACGGTGGCTCTACTGCCCTCCATATGCTACAAATAGCTGGCCACCTAGAG 444  
Db 121 CAAAGTGGGATACGGTGGCTCTCAGGCCCTCCATAGCTGCAATAGCTGGACACCCAGAG 180  
Qy 445 GCTGCTGATGCTGTTGCAACATGAGAGCTAATGTCAATATTCAAGATGCGAGTTTTTTC 504  
Db 181 GCTGCAAGAGTGTGCTTCAACATGGGCCAATGTGAATGTTCAAGATGCGGCTCTTCTTC 240  
Qy 505 ACTCCATTGCGATATTCAGCGCTACTATGGACATGAACAGGTAACTCGGCTTCTTTTGA 564  
Db 241 ACCCCACTGCGACATTGCGAGCTACTATGGCCACGAGCAGGTAAACAGTGTCTCTTTGA 300  
Qy 565 TTTGCTGCTGATGAATAATGTAAGTGGTGAAGTGGAGTAGACCCCTCCACCTAGCATCT 624  
Db 301 TTTGGTGGCTGATGTCAATGTAAAGCGGTGAAGTGGGACAGGCCCTCTGCGACCTGGCCTCT 360  
Qy 625 GCAAAAGGATTTCTTGAATATTGCAAAATCTTTGATGGAAGAGGAGGAGCAAGCAGATGTG 684  
Db 361 GCAAGGGCTTCTTCAACATTGTGAAACTCTCTGTAAGAGGAGGAGCAAGCAGATGTG 420  
Qy 685 AATGCTCAAGATTAATGAAGACCATGTCCACCTCCATTTCTGTTCTCGATTTGGACACCAT 744  
Db 421 AAGCGTCAGACAAATGAAGACCAACGTCCTCTGCACTTCTGTTCTCGATTTGGACACCAT 480  
Qy 745 GATATAGTTAAGTATCTGCTGCAAAAGTGAATTTGGAAGTTCACACCTCATGTTGTTAATATC 804  
Db 481 AATATAGTAGCTTACTGCTCCAGAGTGACTTANAGGTCACGCTCAGCTCATTAACATC 540  
Qy 805 TATGGAGATACCCCTTACACCTGGCATGCTCAATGGCAAAATTTGAAGTT 855  
Db 541 TATGGTGACACTCTTTGCACTGGCATGCTCAATGGAAATTTTGAANTT 591

## RESULT 11

US-09-833-381-1151  
; Sequence 1151, Application US/09833381  
; Patent No. US20020132090A1  
; GENERAL INFORMATION:  
; APPLICANT: Robison, Keith E.  
; TITLE OF INVENTION: No. US20020132090A1e1 Nucleic Acid and Protein Homologs  
; FILE REFERENCE: 5800-119  
; CURRENT APPLICATION NUMBER: US/09/833,381  
; PRIOR FILING DATE: 2001-04-11  
; PRIOR APPLICATION NUMBER: 09/516,448  
; NUMBER OF SEQ ID NOS: 2050  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1151  
; LENGTH: 740  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)...(740)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-833-381-1151

Query Match 16.5%; Score 413.2; DB 9; Length 740;  
Best Local Similarity 79.0%; Pred. No. 1.1e-114;  
Matches 498; Conservative 0; Mismatches 131; Indels 1; Gaps 1;  
Qy 1 ATGGGAAATTAATAATCTAGACCAACCCAAACTTGTACTGATGAATGGAGAAAAGTC 60  
Db 73 ATGGGAAATTAATAATCTAGACCAACCCAAACTTGTCTGATGAATGGAGAAAAGT 132  
Qy 61 AGTGAATCATATGTTATCAATAGAAAGATTAGAAGATGACCTGCAGATCAAGGAAAA 120  
Db 133 AGTGAATCTTACGCTATTATCATAGAAAGCTGGAGGATAAAGCTGCAGATCAAGAAAT 192  
Qy 121 GAAGTGCAGAGAACTAAGGAATATATTTGGCTCTGTATGAAGCCCTCAGTAAAGTCAATTTA 180  
Db 193 GAATTTCAAGAACTAAGGCACATCTTTGGCTCTGTATGAAGCCCTCAGTGAAGTCAGTTTA 252

QY 181 AATTACCGCACTGAAAAAGGGCTGTCTTACTTCTATTTATTTGTCATTTTGTGAGGCAAG 240  
Db 253 AATTACCGCACTGAAAAAGGGCTGTCTTACTTCTATTTATTTGTCATTTTGTGAGGCAAG 312  
QY 241 AATACACATATTCGAACCTTTATGTTGAAGGGCTCGCCCACTCTGACCTGACAGAAAT 300  
Db 313 AAGTACACATATTCGGTGGCCCTTTATGTTTAAAGGGCTCGCTCCATCCAGACTGACGAAAT 372  
QY 301 GGATTTACAGCCTTCATTTAGCAGTTTACAGAGTTTACAGAGTAATGACAGATTTGATCTTCTG 360  
Db 373 GGGTTTACAGCTCTGCACTGCGCGTTTACAGAGACCGCGGACCTTATCATCTTACTG 432  
QY 361 CTTTAC-AGTGAGCTGATATACAGCAGGTTTGGATACGGTGGCTCACTGCCCTCCATAT 419  
Db 433 TTGCAACAGCGGAGCAGATGTTTACAGAGTGGGATACGGTGGCTCACAGCCCTCCACAT 492  
QY 420 TGCTACATAGCTGCGCACCTAGAGGCTGCTGATGCTGTTGCAACATGAGCTAATGT 479  
Db 493 AGCTGCAATAGCTGACACCCAGAGGCTGCANAAAGTGTGNTGCAACATGGGCGCAACGT 552  
QY 480 CAATATTCAAGATGAGCTTTTTCCTCCTCCATTTGCATATTTGCAGCGTACTATGGACATGA 539  
Db 553 GAATGTTCAAGATGCGCTTCTTTCACCCCACTGCATTTGCAGCTTACTATGGCAAGA 612  
QY 540 ACAGTAACTGCGCTTCTTTTGAATTTGGTGTGATGTAATGTAAGTGTGAAAGTTGG 599  
Db 613 GCAGGTAACCAAGTGTCTTTTGAAGTTTGGNGGTGATGTCNATGTAAGCGGTNAAGTTGG 672  
QY 600 ACATAGACCCCTCCACCTAGCATCTGCAAA 629  
Db 673 GNACAGGCTCGGCCCTTGGCCNCGNAAA 702

## RESULT 12

US-09-833-381-1147/c  
; Sequence 1147, Application US/09833381  
; Patent No. US20020132090A1  
; GENERAL INFORMATION:  
; APPLICANT: Robison, Keith E.  
; TITLE OF INVENTION: No. US20020132090A1el Nucleic Acid and Protein Homologs  
; FILE REFERENCE: 5800-119  
; CURRENT APPLICATION NUMBER: US/09/833,381  
; CURRENT FILING DATE: 2001-04-11  
; PRIOR APPLICATION NUMBER: 09/516,448  
; PRIOR FILING DATE: 2000-02-29  
; NUMBER OF SEQ ID NOS: 2050  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1147  
; LENGTH: 304  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(304)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-833-381-1147

Query Match 7.4%; Score 184.2; DB 9; Length 304;  
Best Local Similarity 79.6%; Pred. No. 5.1e-45;  
Matches 242; Conservative 0; Mismatches 59; Indels 3; Gaps 2;  
QY 266 TGAAGGGCTCGGCCATCTCGACTGACAGAAATGGAATTACAGCC--TTGCAATTTAGC 323  
Db 304 TAAAGGGCTCGTCCATCCAGACTGACGAGAAATGGGTTTCCAAAGCTTTGCACTGCC 245  
QY 324 AGTTTACAGGTAATGAGAAATGATCAC--TTCTCTGCTTCAAGTGGAGCTGATATAC 382  
Db 244 GGTTTTACAGGACAGCCCGGAATTTATCATTCTTCACTGTTGACAGCGGAGCAGATGTT 185  
QY 383 AGCAGGTTGGATACGGTGGCTCTCACTGCCCTCCCATTTGCTACAATAGCTGGCCACCTAG 442  
Db 184 AGCAAGTGGGATACGGTGGCTCACAGCCCTCCACATAGCTGCAATAGCTGGACACCCAG 125

QY 443 AGGCTGTGATGTCTGTTTGCACATGGAGCTAATGTCATATTCAGATGCGAGTTTTT 502  
Db 124 AGGCTGCAGAAAGTGTCTGATGCAACATGGGGCCAAACGTGAATGTTCAAGATGCGGTCTTCT 65  
QY 503 TCATCTCATTCATATTTGAGCGTACTATGGACATGAAAGGTAACCTCCGCTTCTTTGA 562  
Db 64 TCACCCCACTGCACATTCGACGNTACTATGGGCACGAGCGGTAACCACTGTCCTTTGA 5  
QY 563 AATT 566  
Db 4 AGTT 1  
RESULT 13  
US-10-424-599-53404  
; Sequence 53404, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 53404  
; LENGTH: 1953  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_19237C.1  
US-10-424-599-53404

Query Match 4.4%; Score 109.6; DB 13; Length 1953;  
Best Local Similarity 51.3%; Pred. No. 1.3e-21;  
Matches 390; Conservative 0; Mismatches 349; Indels 21; Gaps 5;  
QY 1405 ATTGCGTCAGGTCTTTTGGAAAGTATATATAAGGACGATGCGAGAAATATAAATAGTGGT 1464  
Db 884 ATTGCATCTGGTCATATGTTGAACATTTTAAAGGTGTATATTGTAGCCAGGAAGTGGCC 943  
QY 1465 ATAAACGTTATCGAGCCATACCTACTGCTCCAAGTCAGATGTGATATGTTTGGCGA 1524  
Db 944 ATCAAAGTTCTCAAGGCTGACCATGTAATCAAGATTCAGAGAGA---GTTCGCACAG 1000  
QY 1525 GAGGTGTCCATTCTCTGCCAGCTCAATCATCCCTGCGTAATTCAGTTGTGGTGGTCTTGC 1584  
Db 1001 GAAGTCTATATCATGAGAAAGTTTCGACACAAAGATGTTGTACAAATTCATAGGAGCTGT 1060  
QY 1585 TTGAATGATCCGACCCAGTTTGCATGTCATCTCAATACATATCAGGGGTTCTCTGTC 1644  
Db 1061 ACCAAGTCCCAAGCT---TATGCAATGTAACAGAAATTTATGCTGTGGTGAAGTGTGAT 1117  
QY 1645 TCCCTCCTTCATGACGAGAGAGGATCTTGTGATTTTGCAGTCTAAATTAATTTATTCAGTA 1704  
Db 1118 GACTACCTACATAAGCAGAGAGGCTTTTAAATTTTCTTACACTGCTCAAGTAGCAAT 1177  
QY 1705 GATGTTGCCAAAGCATGAGTACCTTCAACACCTGACACGCCAATATACATCGTGAC 1764  
Db 1178 GATGTTTCCAAAGGAATGAATCTACTTGCACCA-----ACATAATATATCATAGAGAC 1231  
QY 1765 TTGAACAGTCACAAATTTCTCTATGAGGATGGGCACTGCTGTGGTGGCAGATTTTGA 1824  
Db 1232 TTGAGGCTGCCAACCTTTTGTAGGATGAAATTTGACTGTAAGGTGCTGATTTTGGG 1291  
QY 1825 GAATCAAGATTTCTACAGTCTCTGGATGAAGACACATGACAAAACACCTGGGAACCTC 1884  
Db 1292 GTTGCTAGATTAAA-----GCTCAATCTGGCTCATGACTGCGAGAAACTGGAACATAT 1345

```
QY 1885 CGTTGGATGCTCTGAGGTTTCACCGAGTGCACCTCGGTACACCATCAAGCAGATGTC 1944
Db 1346 CGATGGATGGCTCTGAGGTTTAGAACA---CAAGCGGTATGATCACAAGCTGATGTA 1402
QY 1945 TTACAGTATGCTCTGTGCTGTGGGAAATTCCTCACTGGGAAATTCATTCGCTCATCTC 2004
Db 1403 TTTAGTTTGGAAATGTTTATGGAGTTGCTCACCAGAAAGCTTCCATACGAATATTTA 1462
QY 2005 AAGCAGCGGCTCGGCAGCAGACATGGCTTACCACCATCAGACCTCCCTATGGCTAT 2064
Db 1463 ACCCCCTACAGGAGCTATAGGAGTGGTTCAAAGGGTTGCGAGCCACCATCCGGAAG 1522
QY 2065 TCATTCGCCAAGCCATPATCATCTCTGCTGATACGAGGGTGAACGCATGCTCTGAAGA 2124
Db 1523 AACACTCATCAAAGTTTGGAGCTCTTGAGAGGCTCTGGCAGCAAGATCCTACATTG 1582
QY 2125 AGACCCGAATTTCTGAAGTTGTCATGAAGTTAGAAGT 2164
Db 1583 AGACCTGATTTCTCGGAATATTCGAGATCCTCGACAGT 1622

RESULT 14
US-10-424-599-129156
; Sequence 129156, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 129156
; LENGTH: 3238
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_87633C.1
US-10-424-599-129156

Query Match 4.3%; Score 107.4; DB 13; Length 3238;
Best Local Similarity 52.4%; Pred. No. 8.e-21;
Matches 387; Conservative 0; Mismatches 331; Indels 21; Gaps 6;

QY 1406 TTGCTCAGGTTCTTTGGGAAGTATATAAGGACGATCGAGATAAATAGTGGCTA 1465
Db 1615 TTGCTCTGGGTCAATTTGGTGCTTTGACAGGTACATATTGTAGTCAAGATGGCTA 1674
QY 1466 TAAACGTTATCGAGGCAATPACCTACTGCTCCAAAGTCAGATGTGGATATGTTTGGCCAG 1525
Db 1675 TCAAA-GTCTTAAGCTGAGCGCATAGTACAGATATGCTGAGG--GAGTTTGCACAGG 1731
QY 1526 AGGTGTCATCTCTGCCACTCAATCATCCCTGCGTAATTCAGTTTGGGTGCTTGT 1585
Db 1732 AAGTTTATATCATGAGGAAGATTCGACACAAGAATGTTGTCAGTTTCATTTGCGCATG-- 1789
QY 1586 TGAATGATCCAGCGAGTTGGCCATGTCTACTCAATACATATCAGGGGTTCTCTGTCT 1645
Db 1790 -TACTAGGCCCCAAATCTTTGTTATGTTACTGATGTTATGCTAGGGAAGCTTATG 1848
QY 1646 CCCTTCCTATGAGCAGAGAGATTTGATTTGGAGTCTAAATTAATTATGAGTAG 1705
Db 1849 ACTTCTGCAACAACAAGAGGTGTATTTAAGCTTCCATCTTTTGCTTAAAGTAGCAATG 1908
QY 1706 ATGTTGCCAAGGCTGAGTACCTTCACAAACCTGACAGGCAATATACATCTGACT 1765
Db 1909 ATGTTTCCAGGGATGAATCTATTGACCA-----AAATATATTAATTCACAGGACC 1962
QY 1766 TGAACAGTCAATATTTCTTCTCTATGAGGATGGGATGCTGTGGCAGATTTTGGAG 1825
```

```
Db 1963 TCAGACTGCCATCTTCTGATGGATGAATGAAGTGTCAAGGTTGCTGATTTGGGG 2022
QY 1826 AATCAAGATTTTACAGTCTCTGGATGAAGACAAACATGACAAAAAACCTGGGAACCTCC 1885
Db 2023 TTGCGAGGTGCAA-----ACTCAGTCTGGAGTGATGACAGCTGAAACTGGAACATACC 2076
QY 1886 GTTGGATGGCTCTGAGGTTTCACGCGAGTGCACCTGGGTACACCATCAAGCAGATGTCT 1945
Db 2077 GTTGGATGGCTCTCTGAGG---TCATTGAACACAACCATATACCAAGAGGAGATGTT 2133
QY 1946 TCAGCTATGCTCTGTCTGTGGGAAATTCCTCACTGGGAAATTCATTCGCTCATCTCA 2005
Db 2134 TCAGTTTGGAAATAGCTCTTTGGAGCTTTTAACTGGAGAACTGCTTACTCTTCCTGGA 2193
QY 2006 AGCAGCGGCTCGGCAGCAGACATGGCTTACCACCATCAGACCTCCCATTTGGCTATT 2065
Db 2194 CCCATTACAGCAGAGTTGGCGTGTGCGAAGGCGCTCGGGCTACAATTCCTCAAAA 2253
QY 2066 CCATTCGCCAAGCCCATATCATCTCTGCTGATACGAGGTGGAACGCATGTCTCTGAAGGAA 2125
Db 2254 ATAGCAGCCCAAGACTTTCTGAACTTCTTCAGCGGTGCTGGCAACAAGATCCAACGCAA 2313
QY 2126 GACCCGAATTTTCTGAAGT 2144
Db 2314 GACCAAAATTTCTCTGAAAT 2332

RESULT 15
US-10-425-114-22955
; Sequence 22955, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 22955
; LENGTH: 2207
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3592-081-B1_FLI
US-10-425-114-22955

Query Match 4.0%; Score 100.2; DB 13; Length 2207;
Best Local Similarity 52.1%; Pred. No. 1.e-18;
Matches 328; Conservative 0; Mismatches 283; Indels 18; Gaps 4;

QY 1516 TTTTGGCGAGAGTGTCATCTCTGCCAGCTCAATCATCCCTCGGTAAATTCAGTTTGTG 1575
Db 1133 TTTGCTCAGGAGTATATATATATGAGGAAGGTCGCTCAACAAGATGTTGTGCAATTTAT 1192
QY 1576 GGTGCTTGTGTAATGATCCAGCCAGTTTGGCATTTGTCACCTCAATACATATCAGGGGT 1635
Db 1193 GGGCTTGGACTTAACCCCAACCTATGT---ATAGTCACAGAAATTTATGTCGGTGA 1249
QY 1636 TCTCTGTCTCCCTCTTTCATGAGCAGAGAGATTTCTTGATTTTGCAGTCTAAATTAAT 1695
Db 1250 AGTGTGTATGATTACCTCCATAAACAAGGTGTTTTCAAACTTCTCTGCTTTAGTTGA 1309
QY 1696 ATTGCAGTAGATCTTCCCAAGCATGGAGTACCTTCACAACCTGCACAGCCCAATTATA 1755
Db 1310 GTTGCAACGGATGTGTCAAAAGGCATGAGCTACTTACACCAAGATAAT-----ATTATT 1363
```

QY	1756	CATCGTGA	CTTGA	CAGTCA	CAATAT	CTCTCT	TATGAG	GATGG	CA	TGCTGT	GGCA	1815
Db	1364	CATCGAT	TTTGA	AACTG	CAAA	CTTCTG	ATGAT	GAAT	GG	ACTG	TTAA	1423
QY	1816	GATTTGG	AGAA	TCTA	AGAT	TTTCTA	CAGT	CTCT	GG	ATGA	GA	1875
Db	1424	GATTTGG	TGTGG	CACG	TGTTAA	AGCTCA	ATCTG	GAGT	-----	ATG	ACTG	1477
QY	1876	GGGA	ACCTCC	GTG	ATGG	CTCCTG	AGTG	TTCA	CG	CAGT	GCAC	1935
Db	1478	GGTACT	TACCG	TGG	ATGG	CCCC	CAGAG	GTCA	TTG	AGCA	---CA	1534
QY	1936	GCAG	ATGTCT	TCAG	CTAT	GTCTGT	GTGG	AAAT	TTCT	CACTG	CG	1995
Db	1535	GCTG	ATGTTT	TAGT	TTTGG	AATTTG	ATGTG	GG	AACT	ACTC	AC	1594
QY	1996	GCTCA	TCTCA	AGCC	AGCG	CTGG	GCAG	CA	CATGG	CTTAC	CA	2055
Db	1595	GAGT	ACCTG	ACTCC	ACTCA	AGCCG	CTGTTG	GTGTG	CGA	AGG	ATTAC	1654
QY	2056	ATTGG	CTATT	CCCA	AGCC	ATAT	CTCT	CTGT	GATAC	GAGG	TGGA	2115
Db	1655	ATCCA	AGCA	TACTC	ACG	CAAG	CTTCC	GAG	TTCT	TCA	GAAT	1714
QY	2116	CCTGA	AGGA	AGCC	CGA	ATTTT	CTG	AGT	2144			
Db	1715	CCTAC	CCAA	AGAC	CAG	ACTTCT	CCG	AAAT	1743			

Search completed: September 6, 2004, 11:52:11  
Job time : 762.842 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 6, 2004, 00:06:13 ; Search time 120.494 Seconds  
(without alignments)  
11537.137 Million cell updates/sec

Title: US-10-626-173-3

Perfect score: 2505

Sequence: 1 atgggaataataaactag.....gcagcttgaggacagcgc 2505

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA:\*  
1: /cgn2\_6/prodata/2/ina/5A COMB.seq:\*  
2: /cgn2\_6/prodata/2/ina/5B COMB.seq:\*  
3: /cgn2\_6/prodata/2/ina/6A COMB.seq:\*  
4: /cgn2\_6/prodata/2/ina/6B COMB.seq:\*  
5: /cgn2\_6/prodata/2/ina/6CTUS COMB.seq:\*  
6: /cgn2\_6/prodata/2/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2505	100.0	2505	3	US-09-291-839-3
2	2505	100.0	2505	4	US-09-458-457-3
3	2505	100.0	2505	4	US-09-947-199A-3
4	2505	100.0	3025	3	US-09-291-839-1
5	2505	100.0	3025	4	US-09-458-457-1
6	2505	100.0	3025	4	US-09-947-199A-1
7	1892.2	75.5	2505	4	US-09-458-457-9
8	1892.2	75.5	2505	4	US-09-947-199A-9
9	1892.2	75.5	3026	4	US-09-458-457-7
10	1892.2	75.5	3026	4	US-09-947-199A-7
11	427.4	17.1	616	4	US-09-833-381-1149
12	414.6	16.6	736	4	US-09-833-381-1150
13	413.2	16.5	740	4	US-09-833-381-1151
14	184.2	7.4	304	4	US-09-833-381-1147
15	84.4	3.4	186	4	US-09-833-381-1148
16	84	3.4	14770	4	US-09-420-132-30
17	78.8	3.1	2894	4	US-09-620-312D-95
18	78.6	3.1	3606	4	US-10-164-595-37
19	78.6	3.1	3454	3	US-09-082-059-1
20	75.4	3.0	4297	4	US-09-972-115A-3
21	75.4	3.0	4493	4	US-09-972-115A-1
22	74.2	3.0	384	4	US-09-372-115A-9
23	74.2	3.0	1069	4	US-09-833-381-841
24	74.2	3.0	3394	4	US-09-427-154-1
25	74.2	3.0	3498	4	US-09-350-982C-4
26	74.2	3.0	3797	4	US-09-696-668-1
27	74.2	3.0	3816	4	US-09-696-668-2

## ALIGNMENTS

### RESULT 1

US-09-291-839-3

; Sequence 3, Application US/09291839A

; Patent No. 6261818

; GENERAL INFORMATION:

; APPLICANT: Raju, Jeyaseelan

; TITLE OF INVENTION: NOVEL CARP PROTEIN AND NUCLEIC ACID MOLECULES AND USES THEREFOR

; FILE REFERENCE: MNI-068

; CURRENT APPLICATION NUMBER: US/09/291,839A

; CURRENT FILING DATE: 1999-04-14

; NUMBER OF SEQ ID NOS: 6

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 3

; LENGTH: 2505

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (1)..(2505)

US-09-291-839-3

Query Match		100.0%	Score 2505;	DB 3;	Length 2505;
Best Local Similarity		100.0%	Pred. No. 0;		
Matches 2505;	Conservative	0;	Mismatches	0;	Gaps
Qy	1	ATGGGAAATTATAAATCTAGACCAACCCAACTTGTACTGATGAATGGAAGAAAAGTC	60		
Db	1	ATGGGAAATTATAAATCTAGACCAACCCAACTTGTACTGATGAATGGAAGAAAAGTC	60		
Qy	61	AGTGAATCATATGTTATCATCAATAGAAAGATTAGAGATGACCTGCAGATCAAGGAAAA	120		
Db	61	AGTGAATCATATGTTATCATCAATAGAAAGATTAGAGATGACCTGCAGATCAAGGAAAA	120		
Qy	121	GAATCAGACGAACCTAAGGAATATATTTGGCTCTGATGAAGCCTTCAGTAAAGTCAATTTA	180		
Db	121	GAATCAGACGAACCTAAGGAATATATTTGGCTCTGATGAAGCCTTCAGTAAAGTCAATTTA	180		
Qy	181	AATTACCGCACTGAATATGGCTGTCTTACTTCTTATTTATGTTGATTTGGAGGCAAG	240		
Db	181	AATTACCGCACTGAATATGGCTGTCTTACTTCTTATTTATGTTGATTTGGAGGCAAG	240		
Qy	241	AAATCATATTCGAACCTCTTATGTTGAAGGGCTCCGCCCCTTCGACTCACAAGAAAT	300		
Db	241	AAATCATATTCGAACCTCTTATGTTGAAGGGCTCCGCCCCTTCGACTCACAAGAAAT	300		
Qy	301	GGATTTACAGCCTTGCAATTTAGCAATTTACAGGAATATGCAAGATTTGATCATTCTCTG	360		
Db	301	GGATTTACAGCCTTGCAATTTAGCAATTTACAGGAATATGCAAGATTTGATCATTCTCTG	360		
Qy	361	CTTCACAGTGGAGCTGATATACAGCAGGTTGGATACGGTGGCTCCTCAGCTCCCTCATATT	420		

Db 361 CTTTCAGTGGAGCTGATATACAGAGGTTGGATACGGTGGCTCTCACTGCCCTCCATATT 420  
Qy 421 GCTACAATAGCTGGCCACTAGAGGCTGCTGATGCTGTGTGAAACATGAGGCTAATGTC 480  
Db 421 GCTACAATAGCTGGCCACTAGAGGCTGCTGATGCTGTGTGAAACATGAGGCTAATGTC 480  
Qy 481 AATATTCAAGATGCAAGTTTTTTTCACTCCATTTGCATATTGACGGTACTATGACATGAA 540  
Db 481 AATATTCAAGATGCAAGTTTTTTTCACTCCATTTGCATATTGACGGTACTATGACATGAA 540  
Qy 541 CAGGTAACCTGCTCTTTTGAATTTGGTCTGATGTAATGTAAGTGGTGAAGTTGGA 600  
Db 541 CAGGTAACCTGCTCTTTTGAATTTGGTCTGATGTAATGTAAGTGGTGAAGTTGGA 600  
Qy 601 GATAGACCCCTCCACTAGCATCTGCAAAAGGATTTTGAATATTGCAAACTCTTGATG 660  
Db 601 GATAGACCCCTCCACTAGCATCTGCAAAAGGATTTTGAATATTGCAAACTCTTGATG 660  
Qy 661 GAAGAAGGAGCAAGAGATGTGAATGCTCAAGATTAATGAAGACCATGTCCTCCAT 720  
Db 661 GAAGAAGGAGCAAGAGATGTGAATGCTCAAGATTAATGAAGACCATGTCCTCCAT 720  
Qy 721 TTCTGTTCTCGAATTTGGACACCATGATATAGTAAAGTATCTGCTGCAAGTGAATTTGGA 780  
Db 721 TTCTGTTCTCGAATTTGGACACCATGATATAGTAAAGTATCTGCTGCAAGTGAATTTGGA 780  
Qy 781 GTTCAACCTCATGTTTAAATATCTATGAGATACCCCTTACACCTGGCATGCTACAAT 840  
Db 781 GTTCAACCTCATGTTTAAATATCTATGAGATACCCCTTACACCTGGCATGCTACAAT 840  
Qy 841 GCAAAATTTGAAGTTGCAAGCAAGCAATCAATCAATATCAGGAACAGAACTCTGACTAAG 900  
Db 841 GCAAAATTTGAAGTTGCAAGCAAGCAATCAATCAATATCAGGAACAGAACTCTGACTAAG 900  
Qy 901 GAAAAATCTTTCAGTGAACAGCTTTTTCATAGTCTGTTGATCTATGCAAGACATTTGAC 960  
Db 901 GAAAAATCTTTCAGTGAACAGCTTTTTCATAGTCTGTTGATCTATGCAAGACATTTGAC 960  
Qy 961 CTAGTCAAAATTTCTTCTGATCAGAAATGTCAATAACATCAACACCAAGGAGGATGG 1020  
Db 961 CTAGTCAAAATTTCTTCTGATCAGAAATGTCAATAACATCAACACCAAGGAGGATGG 1020  
Qy 1021 CACACTGGATTACACTCTGCTGCTACCAAGGTCACATTCGCTGGTTCAAGTTCTTACTG 1080  
Db 1021 CACACTGGATTACACTCTGCTGCTACCAAGGTCACATTCGCTGGTTCAAGTTCTTACTG 1080  
Qy 1081 GATAATGAGCTGATGATGAATCTAGTGGCTTTGATCCAGAGGCTCTAGTGGTGAATAA 1140  
Db 1081 GATAATGAGCTGATGATGAATCTAGTGGCTTTGATCCAGAGGCTCTAGTGGTGAATAA 1140  
Qy 1141 GATGAGCAGACATGTTTCATGCTGCTTATGAAAAGGATGATGCTTGTGACACTC 1200  
Db 1141 GATGAGCAGACATGTTTCATGCTGCTTATGAAAAGGATGATGCTTGTGACACTC 1200  
Qy 1201 CTGAAGACATTAAGAGACCAAGATGAATTTGCCCTGTAATGAATATCTCAGCCTGGA 1260  
Db 1201 CTGAAGACATTAAGAGACCAAGATGAATTTGCCCTGTAATGAATATCTCAGCCTGGA 1260  
Qy 1261 GAGATGCTCTATGCTGCTGCTTCCATCCCTTGGGAGATTAAGACATGACAAA 1320  
Db 1261 GAGATGCTCTATGCTGCTGCTTCCATCCCTTGGGAGATTAAGACATGACAAA 1320  
Qy 1321 GAGAAAGGAGATATTCTCTCTCAAGAGCTGGATTTGCCCTTACATTTTCCATCTTCAGCTC 1380  
Db 1321 GAGAAAGGAGATATTCTCTCTCAAGAGCTGGATTTGCCCTTACATTTTCCATCTTCAGCTC 1380  
Qy 1381 TCAGAAAATTGAGTTCATGATATTGGCTCAGGTTCTTTGGAAAGTATATAAGGA 1440  
Db 1381 TCAGAAAATTGAGTTCATGATATTGGCTCAGGTTCTTTGGAAAGTATATAAGGA 1440  
Qy 1441 CGATGCAAAATAAATAGTGGCTATAAAAGTTATCGAGCCATACCTACTGCTCCAAG 1500

Db 1441 CGATGCAAAATAAATAGTGGCTATAAAAGTTATCGAGCCATACCTACTGCTCCAAG 1500  
Qy 1501 TCAGATGTGATATGTTTTTCCGAGAGGTGTCCATTTCTGCGAGCTCAATCATCCCTGC 1560  
Db 1501 TCAGATGTGATATGTTTTTCCGAGAGGTGTCCATTTCTGCGAGCTCAATCATCCCTGC 1560  
Qy 1561 GTAAATTCAGTTTGGGTGCTTCTGTAATGATCCAGACAGTTTGGCAATCTCACTCAA 1620  
Db 1561 GTAAATTCAGTTTGGGTGCTTCTGTAATGATCCAGACAGTTTGGCAATCTCACTCAA 1620  
Qy 1621 TACATATCAGGGGTTCTCTGTTCTCCCTCTTCATGAGCAGAGAAGGATTTCTGATTG 1680  
Db 1621 TACATATCAGGGGTTCTCTGTTCTCCCTCTTCATGAGCAGAGAAGGATTTCTGATTG 1680  
Qy 1681 CAGTCTAAATTAATTTAGTAGATGTTGCCAAAGGATGGAGTACCTTCACAACCTG 1740  
Db 1681 CAGTCTAAATTAATTTAGTAGATGTTGCCAAAGGATGGAGTACCTTCACAACCTG 1740  
Qy 1741 ACACAGCAATTAATACATCGTAGATTGAACAGTCAATATTCTCTCTATGAGATGG 1800  
Db 1741 ACACAGCAATTAATACATCGTAGATTGAACAGTCAATATTCTCTCTATGAGATGG 1800  
Qy 1801 CATCTGTGTGGCAGATTTTGGAGATCAAGATTTTCAAGTCTCTGATGAAAGCAAC 1860  
Db 1801 CATCTGTGTGGCAGATTTTGGAGATCAAGATTTTCTACAGTCTCTGATGAAAGCAAC 1860  
Qy 1861 ATGCAAAAACAACCTGGACCTCCGTTGGATGGCTCCTGAGGTGTTACGAGTGCCT 1920  
Db 1861 ATGCAAAAACAACCTGGACCTCCGTTGGATGGCTCCTGAGGTGTTACGAGTGCCT 1920  
Qy 1921 CGGTACACCATCAAGCAGATGTTTACAGTATGCTCTGTGTCTGTGGAAATTTCTCACT 1980  
Db 1921 CGGTACACCATCAAGCAGATGTTTACAGTATGCTCTGTGTCTGTGGAAATTTCTCACT 1980  
Qy 1981 GGCAGAAATTCATTCGCTCATCTCAAGCAGGGCTCGGACAGACACATGCTTACCAAC 2040  
Db 1981 GGCAGAAATTCATTCGCTCATCTCAAGCAGGGCTCGGACAGACACATGCTTACCAAC 2040  
Qy 2041 CACATCAGACTCCCATTTGGCTATTCCATTTCCCAAGCCCATATCATCTCTGATAGGA 2100  
Db 2041 CACATCAGACTCCCATTTGGCTATTCCATTTCCCAAGCCCATATCATCTCTGATAGGA 2100  
Qy 2101 GGGTGAACGATGCTCTGAAGGAAGACCCGAATTTTCTGAAGTTGTCTAAGTTAGNA 2160  
Db 2101 GGGTGAACGATGCTCTGAAGGAAGACCCGAATTTTCTGAAGTTGTCTAAGTTAGNA 2160  
Qy 2161 GAGTGTCTCTGCAACATTTGAGCTGATGTTCTCTGATCAAGTAAAGAGTGGTCTCTC 2220  
Db 2161 GAGTGTCTCTGCAACATTTGAGCTGATGTTCTCTGATCAAGTAAAGAGTGGTCTCTC 2220  
Qy 2221 TCACCTTCTTCTTCTGATTCCTGTAACCCGGAGGACCTGGCCGAGTCAATG 2280  
Db 2221 TCACCTTCTTCTTCTGATTCCTGTAACCCGGAGGACCTGGCCGAGTCAATG 2280  
Qy 2281 GCAGCATTAAGAGTCTGTTTGAATTTGGAATATGCTCTAAATGCAAGTCTTATGCTCT 2340  
Db 2281 GCAGCATTAAGAGTCTGTTTGAATTTGGAATATGCTCTAAATGCAAGTCTTATGCTCT 2340  
Qy 2341 TTGTCCCAAAAGTCTGGACAAATATTCTCTCAAGGTCTGTCTTTGGAGAGATGAAAGA 2400  
Db 2341 TTGTCCCAAAAGTCTGGACAAATATTCTCTCAAGGTCTGTCTTTGGAGAGATGAAAGA 2400  
Qy 2401 AGTCTTCAATACACCCATTAAGATGATGGTATGATCCGATCCCATGAGCTCAATG 2460  
Db 2401 AGTCTTCAATACACCCATTAAGATGATGGTATGATCCGATCCCATGAGCTCAATG 2460  
Qy 2461 CATTTTCAATTTCTTCCGAAATAGTAGCAGCTTTTGGAGACAGCAGC 2505  
Db 2461 CATTTTCAATTTCTTCCGAAATAGTAGCAGCTTTTGGAGACAGCAGC 2505

```
; Sequence 3, Application US/09458457
; Patent No. 6500654
; GENERAL INFORMATION:
; APPLICANT: Raju, Jeyaseelan
; TITLE OF INVENTION: NOVEL DARK PROTEIN AND NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: mni-068cp
; CURRENT APPLICATION NUMBER: US/09/458,457
; EARLIER FILING DATE: 1999-12-10
; EARLIER APPLICATION NUMBER: 60/111,938
; EARLIER FILING DATE: 1998-12-11
; EARLIER APPLICATION NUMBER: 09/291,839
; EARLIER FILING DATE: 1999-04-14
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2505
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2505)
US-09-458-457-3

Query Match      100.0%; Score 2505; DB 4; Length 2505;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  ATGGGAATTAATAAATCTAGACCAACCCAAACTGTGACTGATGAATGGAGAAAAAGTC 60
DB      1  ATGGGAATTAATAAATCTAGACCAACCCAAACTGTGACTGATGAATGGAGAAAAAGTC 60

QY      61  AGTGAATCATATGTTATCAATAGAAAAGATTAGAAAGATGACCTGCGAGATCAAGAAAAA 120
DB      61  AGTGAATCATATGTTATCAATAGAAAAGATTAGAAAGATGACCTGCGAGATCAAGAAAAA 120

QY      121  GAATGACAGAACTAGAGGATATATTGGCTCTGATGAGCCCTCAGTAAAGTCAATTTA 180
DB      121  GAATGACAGAACTAGAGGATATATTGGCTCTGATGAGCCCTCAGTAAAGTCAATTTA 180

QY      181  AATTACCGCACTGAAATGGGCTGTCTCTACTTCATTATTTGTCATTTGTCAGGGCAAG 240
DB      181  AATTACCGCACTGAAATGGGCTGTCTCTACTTCATTATTTGTCATTTGTCAGGGCAAG 240

QY      241  AATACATATTTGAACTCTTATTTGAAAGGGCTCCGCCATCTCGACTGACAGAAAT 300
DB      241  AATACATATTTGAACTCTTATTTGAAAGGGCTCCGCCATCTCGACTGACAGAAAT 300

QY      301  GGATTTACAGCTTGCATTTAGCAGTTTACAAGGATAATGCGAATTTGATCACTTCTCTG 360
DB      301  GGATTTACAGCTTGCATTTAGCAGTTTACAAGGATAATGCGAATTTGATCACTTCTCTG 360

QY      361  CTTACAGTGGAGCTGATATACAGAGGTTGGATACGGTGGCTCAGTCCCTCCATATT 420
DB      361  CTTACAGTGGAGCTGATATACAGAGGTTGGATACGGTGGCTCAGTCCCTCCATATT 420

QY      421  GCTACATATCTGCCACCTAGAGGCTGTGATGCTGTCTGTCACATGGAGCTAATGTC 480
DB      421  GCTACATATCTGCCACCTAGAGGCTGTGATGCTGTCTGTCACATGGAGCTAATGTC 480

QY      481  AATATTCAGAGTCAAGTTTTTTTCACTCCATTTGATTAATTCAGCGTACTATGGACATGAA 540
DB      481  AATATTCAGAGTCAAGTTTTTTTCACTCCATTTGATTAATTCAGCGTACTATGGACATGAA 540

QY      541  CAGGTAATCTGCCCTCTTTTGAATTTGGTCTGATGTAATTAAGTGGTGAAGTTGGA 600
DB      541  CAGGTAATCTGCCCTCTTTTGAATTTGGTCTGATGTAATTAAGTGGTGAAGTTGGA 600

QY      601  GATAGACCCCTCCACCTAGCATCTGCAAAAGAGATTTCTTGAATATTGCAAAAATCTTTGATG 660
DB      601  GATAGACCCCTCCACCTAGCATCTGCAAAAGAGATTTCTTGAATATTGCAAAAATCTTTGATG 660

QY      661  GAAGAAGGACAAAGCAGATGTAATGTCTCAAGATTAATGAAGACCATGTGCCACTCCAT 720
```

```
DB      661  GAAGAAGGACAAAGCAGATGTAATGTCTCAAGATTAATGAAGACCATGTGCCACTCCAT 720
QY      721  TTCTGTCTTCGATTTGGACACCATGATATAGTTAGTATCTCTGCTGCAAAAGTGAATTTGAA 780
DB      721  TTCTGTCTTCGATTTGGACACCATGATATAGTTAGTATCTCTGCTGCAAAAGTGAATTTGAA 780
QY      781  GTTCAACCTCATGTTGTTAATATCTATGAGAGTATCCCCCTTACACCTGGCATGCTACAAT 840
DB      781  GTTCAACCTCATGTTGTTAATATCTATGAGAGTATCCCCCTTACACCTGGCATGCTACAAT 840
QY      841  GGCAAATTTGAAGTTGCCAAGGAAATCATCCAAATATCAGGAACAGAAAGTCTGACTAAG 900
DB      841  GGCAAATTTGAAGTTGCCAAGGAAATCATCCAAATATCAGGAACAGAAAGTCTGACTAAG 900
QY      901  GAAAAATCTCTCAGTGAAACAGCTTTTCACTAGTGTGTGATGCTATGGCAAGAGCATTCAC 960
DB      901  GAAAAATCTCTCAGTGAAACAGCTTTTCACTAGTGTGTGATGCTATGGCAAGAGCATTCAC 960
QY      961  CTAGTCAAAATTTCTTCTTGATCAGAAATGTCTATAAATCAACCCAAAGGAGGATGGG 1020
DB      961  CTAGTCAAAATTTCTTCTTGATCAGAAATGTCTATAAATCAACCCAAAGGAGGATGGG 1020
QY      1021  CACTGGATTACACTCTGCTTGTACACGGTCAATTCGGCTGTTCAAGTCTTACTG 1080
DB      1021  CACTGGATTACACTCTGCTTGTACACGGTCAATTCGGCTGTTCAAGTCTTACTG 1080
QY      1081  GATAATGGAGCTGATATGAATCTAGTGGCTTGTGATCCAGCAGGTCTAGTGGTGAATAA 1140
DB      1081  GATAATGGAGCTGATATGAATCTAGTGGCTTGTGATCCAGCAGGTCTAGTGGTGAATAA 1140
QY      1141  GATGAGCAGACATGTTGATGTTGGCTTATGAAAAGGSCATGATGCCATTTGTCACATC 1200
DB      1141  GATGAGCAGACATGTTGATGTTGGCTTATGAAAAGGSCATGATGCCATTTGTCACATC 1200
QY      1201  CTGAAGCATTTATAAGAGACCAAGATGAATTCCTGCTGTAATGAATATTTCTCAGCCTGGA 1260
DB      1201  CTGAAGCATTTATAAGAGACCAAGATGAATTCCTGCTGTAATGAATATTTCTCAGCCTGGA 1260
QY      1261  GGAGATGGCTCTTATGCTGTCTTCCATCACCCTTGGGGAAGATTAAAGCATGACAAAA 1320
DB      1261  GGAGATGGCTCTTATGCTGTCTTCCATCACCCTTGGGGAAGATTAAAGCATGACAAAA 1320
QY      1321  GAGAAGCAGATATTTCTCTTACAGCTGATGCTTCCATTTCCATTTCCATTTCCATTTCCAT 1380
DB      1321  GAGAAGCAGATATTTCTCTTACAGCTGATGCTTCCATTTCCATTTCCATTTCCATTTCCAT 1380
QY      1381  TCAGAAATTTGAGTTCCATGAGATTTAGGCTCAGGTTCTTTTGGGAAAGTATATAAGGA 1440
DB      1381  TCAGAAATTTGAGTTCCATGAGATTTAGGCTCAGGTTCTTTTGGGAAAGTATATAAGGA 1440
QY      1441  CGATGAGAAATTAATAATAGTGGCTATAAAACGTTATCGAGCCCAATACCTACTGCTCCAAG 1500
DB      1441  CGATGAGAAATTAATAATAGTGGCTATAAAACGTTATCGAGCCCAATACCTACTGCTCCAAG 1500
QY      1501  TCAGATGTGATATGTTTTCGAGAGGTTGCCATTTCTGCGAGCTCAATCATCCCTGC 1560
DB      1501  TCAGATGTGATATGTTTTCGAGAGGTTGCCATTTCTGCGAGCTCAATCATCCCTGC 1560
QY      1561  GTAATTCAGTTTGTGGGTGCTTGTGTTCCCTTCTGATGAGCAAGGATTTGCCATTTGCTCAA 1620
DB      1561  GTAATTCAGTTTGTGGGTGCTTGTGTTCCCTTCTGATGAGCAAGGATTTGCCATTTGCTCAA 1620
QY      1621  TACATATCAGGGGTTCTCTGTTCTCCCTTCTGATGAGCAAGGATTTCTTGAATTTG 1680
DB      1621  TACATATCAGGGGTTCTCTGTTCTCCCTTCTGATGAGCAAGGATTTCTTGAATTTG 1680
QY      1681  CAGTCTAAATTAATTTATTCAGTAGATGTTGCCAAAGCATGGAGTACCTTCAACAACCTG 1740
DB      1681  CAGTCTAAATTAATTTATTCAGTAGATGTTGCCAAAGCATGGAGTACCTTCAACAACCTG 1740
QY      1741  ACACAGCCATTTATACATGCTGACTTGACATGCTCAATATTTCTTCTATGAGGATGGG 1800
```

Db 1741 ACACAGCCAAATATACATCGTGACTTGAACAGTCACAAATATCTTCTCTATGAGGATGG 1800  
QY 1801 CATGCTGTGGGCGAGATTTGGAGAAATCAAGATTTCTACAGTCTCTGGATGAAGCAAC 1860  
Db 1801 CATGCTGTGGGCGAGATTTGGAGAAATCAAGATTTCTACAGTCTCTGGATGAAGCAAC 1860  
QY 1861 ATGACAAAACCACTGGGAACCTCGTTGGATGCTCTGAGGTGTTTCAGCAGTGCAC 1920  
Db 1861 ATGACAAAACCACTGGGAACCTCGTTGGATGCTCTGAGGTGTTTCAGCAGTGCAC 1920  
QY 1921 CGGTACACCAATCAAAAGCAGATGTTCTTCAGCTATGCTCTGTGTGTGGGAAATTCAC 1980  
Db 1921 CGGTACACCAATCAAAAGCAGATGTTCTTCAGCTATGCTCTGTGTGTGGGAAATTCAC 1980  
QY 1981 GGGGAATTCATTCGCTCATCTCAAGCCAGCGGCTGGGAGAGATGGCTTACCAC 2040  
Db 1981 GGGGAATTCATTCGCTCATCTCAAGCCAGCGGCTGGGAGAGATGGCTTACCAC 2040  
QY 2041 CACATCAGACCTCCCATTTGGCTATTTCCATTTCCCAAGCCCATATCATCTCTGCTGATACGA 2100  
Db 2041 CACATCAGACCTCCCATTTGGCTATTTCCATTTCCCAAGCCCATATCATCTCTGCTGATACGA 2100  
QY 2101 GGGTGAACGCAATGCTCGTAGGAAGACCCGAATTTCTGAAAGTTGTCATGAAGTTAGAA 2160  
Db 2101 GGGTGAACGCAATGCTCGTAGGAAGACCCGAATTTCTGAAAGTTGTCATGAAGTTAGAA 2160  
QY 2161 GAGTGTCTCTGCAACATTTGAGTGTCTCTGATGCTCTGATCAAGTACAGCAGTGGTCTCTC 2220  
Db 2161 GAGTGTCTCTGCAACATTTGAGTGTCTCTGATGCTCTGATCAAGTACAGCAGTGGTCTCTC 2220  
QY 2221 TCACCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2280  
Db 2221 TCACCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2280  
QY 2281 GCAGCATTAAGAAAGTCGTTTCGAATTTGAATATGCTCTAAATGCAAGGTCCTATGCTGCT 2340  
Db 2281 GCAGCATTAAGAAAGTCGTTTCGAATTTGAATATGCTCTAAATGCAAGGTCCTATGCTGCT 2340  
QY 2341 TTGTCCCAAGTGTGGAACAATATTTCTCTCAAGGTCTGTCTTTTGGAGGAGATGAAAAGA 2400  
Db 2341 TTGTCCCAAGTGTGGAACAATATTTCTCTCAAGGTCTGTCTTTTGGAGGAGATGAAAAGA 2400  
QY 2401 AGTCTTCAATACACACCAATGACAAATATGCTATGATCCGATCCATGAGTCAATG 2460  
Db 2401 AGTCTTCAATACACACCAATGACAAATATGCTATGATCCGATCCATGAGTCAATG 2460  
QY 2461 CATTTTCATTTCTGCGGAAATAGTAGCAGCTTTGAGGACAGCAGC 2505  
Db 2461 CATTTTCATTTCTGCGGAAATAGTAGCAGCTTTGAGGACAGCAGC 2505

## RESULT 3

US-09-947-199A-3  
; Sequence 3, Application US/09947199A  
; Patent No. 6660490  
; GENERAL INFORMATION:  
; APPLICANT: Raju, Jeyaseelan  
; TITLE OF INVENTION: NOVEL CARB PROTEIN AND NUCLEIC ACID MOLECULES AND USES  
; FILE OF INVENTION: THEREFOR  
; FILE REFERENCE: MNI-068CP2  
; CURRENT APPLICATION NUMBER: US/09/947,199A  
; CURRENT FILING DATE: 2001-09-05  
; PRIOR APPLICATION NUMBER: US 60/111,938  
; PRIOR FILING DATE: 1998-12-11  
; PRIOR APPLICATION NUMBER: US 09/291,839  
; PRIOR FILING DATE: 1999-04-14  
; PRIOR APPLICATION NUMBER: US 09/458,457  
; PRIOR FILING DATE: 1999-12-10  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 2505  
; TYPE: DNA

; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(2505)  
US-09-947-199A-3

Query Match 100.0%; Score 2505; DB 4; Length 2505;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGAATTTATAATCTAGACCAACCCAAACTTGTACTGATGAGTGGAGAAAAGTC 60  
Db 1 ATGGGAATTTATAAATCTAGACCAACCCAAACTTGTACTGATGAGTGGAGAAAAGTC 60  
QY 61 AGTCAATCATATGTTATCACAATAGAAAGATTAGAAAGATGACCTGCAGATCAAGAAAA 120  
Db 61 AGTCAATCATATGTTATCACAATAGAAAGATTAGAAAGATGACCTGCAGATCAAGAAAA 120  
QY 121 GAACTGACAGAACTAAGGAATATATTTGGCTCTGATGAAGCCTTCAGTAAAGTCAATT 180  
Db 121 GAACTGACAGAACTAAGGAATATATTTGGCTCTGATGAAGCCTTCAGTAAAGTCAATT 180  
QY 181 AATTACCGCACTGAAATGGCTCTCTACTTCTATTTATTTGATTTGCGAGGCAAG 240  
Db 181 AATTACCGCACTGAAATGGCTCTCTACTTCTATTTATTTGATTTGCGAGGCAAG 240  
QY 241 AAATCACATATTCGAACTCTTTATGTTGAAAGGGCTCCGCCATCTCGACTGACAAAGAA 300  
Db 241 AAATCACATATTCGAACTCTTTATGTTGAAAGGGCTCCGCCATCTCGACTGACAAAGAA 300  
QY 301 GGATTTACAGCTTGCATTTTAGCAGTTTACAGGATAATACAGATTGATGATCTCTCTG 360  
Db 301 GGATTTACAGCTTGCATTTTAGCAGTTTACAGGATAATACAGATTGATGATCTCTCTG 360  
QY 361 CTTCAAGTGTGAGCTGATATACAGCAGGTTGAGTACGGTGGCTCACTGCCCTCCATATT 420  
Db 361 CTTCAAGTGTGAGCTGATATACAGCAGGTTGAGTACGGTGGCTCACTGCCCTCCATATT 420  
QY 421 GCTCAATAGTGTGCCACCTAGAGGCTGTGATGCTGTGCAACATGGAGTCAATGTC 480  
Db 421 GCTCAATAGTGTGCCACCTAGAGGCTGTGATGCTGTGCAACATGGAGTCAATGTC 480  
QY 481 AATATTCAAGATGCAGTTTTTTTCACTCCATTCATATTCATATTCAGAGATTGATGATG 540  
Db 481 AATATTCAAGATGCAGTTTTTTTCACTCCATTCATATTCATATTCAGAGATTGATGATG 540  
QY 541 CAGTAACTCGCTTCTTTTGAATTTGGTGTGATGTAATTAAGTGGTGAAGTTGA 600  
Db 541 CAGTAACTCGCTTCTTTTGAATTTGGTGTGATGTAATTAAGTGGTGAAGTTGA 600  
QY 601 GATAGACCTTCCACCTAGCATCTGCAAAAGGATTCTTGAATATTCAAAATCTTGTATG 660  
Db 601 GATAGACCTTCCACCTAGCATCTGCAAAAGGATTCTTGAATATTCAAAATCTTGTATG 660  
QY 661 GAAGAAGCAGCAAGCAGATGTAATTCCTCAAGATAATGAAGACCATGTCCCACTCCAT 720  
Db 661 GAAGAAGCAGCAAGCAGATGTAATTCCTCAAGATAATGAAGACCATGTCCCACTCCAT 720  
QY 721 TTCTGTTCTGATTTGACACCATGATATAGTTAAGTATCTGTCGAAAGTCAATTGGAA 780  
Db 721 TTCTGTTCTGATTTGACACCATGATATAGTTAAGTATCTGTCGAAAGTCAATTGGAA 780  
QY 781 GTTCAACTCATGTTGTTAATATCTATGGAGATACCCCTTACACCTGCGCTGCTACAT 840  
Db 781 GTTCAACTCATGTTGTTAATATCTATGGAGATACCCCTTACACCTGCGCTGCTACAT 840  
QY 841 GGCAGAAATTTGAAGTTGCCAAGGAAATCATCCAAATATCAGGAACAGAAATCTGATAG 900  
Db 841 GGCAGAAATTTGAAGTTGCCAAGGAAATCATCCAAATATCAGGAACAGAAATCTGATAG 900  
QY 901 GAAAACATCTTCAGTGAACAGCTTTTTCATAGTGTCTTGTACCTATGCGAAGAGCATTTGAC 960  
Db 901 GAAAACATCTTCAGTGAACAGCTTTTTCATAGTGTCTTGTACCTATGCGAAGAGCATTTGAC 960

QY 961 CTAGTCAAAATTTCTTTGATCAGAAATGCTATTAACATCAACCAACCAAGGAGTGG 1020  
Db 961 CTAGTCAAAATTTCTTTGATCAGAAATGCTATTAACATCAACCAACCAAGGAGTGG 1020  
QY 1021 CACACTGGAATTAACACTCTGCTTGTACCAAGGTGCACATTCGCTGTTTCACTTCTTACTG 1080  
Db 1021 CACACTGGAATTAACACTCTGCTTGTACCAAGGTGCACATTCGCTGTTTCACTTCTTACTG 1080  
QY 1081 GATATGAGCTGATATGAATCTAGTGGCTTGTGATCCCAAGCAGTCTAGTGTGAAAAA 1140  
Db 1081 GATATGAGCTGATATGAATCTAGTGGCTTGTGATCCCAAGCAGTCTAGTGTGAAAAA 1140  
QY 1141 GATGAGCAGACATGTTTGTGATGCGCTTATGAAAAAGGCATGATGCCATTGTCACTC 1200  
Db 1141 GATGAGCAGACATGTTTGTGATGCGCTTATGAAAAAGGCATGATGCCATTGTCACTC 1200  
QY 1201 CTGAAGCATTTAAGAGACCAAGATGAATGTCCTGTGAATGAATTTCTCAGCCTGGA 1260  
Db 1201 CTGAAGCATTTAAGAGACCAAGATGAATGTCCTGTGAATGAATTTCTCAGCCTGGA 1260  
QY 1261 GGAGATGGCTCTATGCTGCTTCCATCACCCTTGGGGAAGATTAAAGCATGACAAA 1320  
Db 1261 GGAGATGGCTCTATGCTGCTTCCATCACCCTTGGGGAAGATTAAAGCATGACAAA 1320  
QY 1321 GAGAAGCAGATATTCCTCTTAAGAGCTGGATGCTTCACTTCCATCTTCACTC 1380  
Db 1321 GAGAAGCAGATATTCCTCTTAAGAGCTGGATGCTTCACTTCCATCTTCACTC 1380  
QY 1381 TCAGAAATGAGTCCATGAGATTAATGCTGCTGCTGCTTCTTGGGAAGTATATAAAGA 1440  
Db 1381 TCAGAAATGAGTCCATGAGATTAATGCTGCTGCTGCTTCTTGGGAAGTATATAAAGA 1440  
QY 1441 CGATGCAAAATAAATAGTGGCTATATAACGGTTATCGAGCCCAATACCTACTGCTCCAA 1500  
Db 1441 CGATGCAAAATAAATAGTGGCTATATAACGGTTATCGAGCCCAATACCTACTGCTCCAA 1500  
QY 1501 TCAGATGATATGTTTTCGCGAGAGTGTCAATCTCTGCGAGCTCAATCACTCCCTGC 1560  
Db 1501 TCAGATGATATGTTTTCGCGAGAGTGTCAATCTCTGCGAGCTCAATCACTCCCTGC 1560  
QY 1561 GTAATTCAGTTTGGGTGCTTCTCAATGATCCAGCAGTTTCCCATTTGTCACCA 1620  
Db 1561 GTAATTCAGTTTGGGTGCTTCTCAATGATCCAGCAGTTTCCCATTTGTCACCA 1620  
QY 1621 TACATATCAGGGGTTCTCTGTTCTCCTCTCATGAGCAGAGGATTTCTGATTTG 1680  
Db 1621 TACATATCAGGGGTTCTCTGTTCTCCTCTCATGAGCAGAGGATTTCTGATTTG 1680  
QY 1681 CAGTCTAAATTAATTTGAGTAGATGTTGCCAAGGATGAGTACCTTCAACACCTG 1740  
Db 1681 CAGTCTAAATTAATTTGAGTAGATGTTGCCAAGGATGAGTACCTTCAACACCTG 1740  
QY 1741 ACACAGCAATTAATATCATGCTGATGAAACAGTCACAATATTTCTCTATGAGGATGG 1800  
Db 1741 ACACAGCAATTAATATCATGCTGATGAAACAGTCACAATATTTCTCTATGAGGATGG 1800  
QY 1801 CATGCTGTGTGGCAGATTTTGGAGATCAAGATTTCTACAGTCTCTGATGAGACAA 1860  
Db 1801 CATGCTGTGTGGCAGATTTTGGAGATCAAGATTTCTACAGTCTCTGATGAGACAA 1860  
QY 1861 ATGACAAAACAACTGGGAACCTCCGTTGGATGGCTCTGAGTGTTCACGAGTGCAT 1920  
Db 1861 ATGACAAAACAACTGGGAACCTCCGTTGGATGGCTCTGAGTGTTCACGAGTGCAT 1920  
QY 1921 CGGTACACCATCAAGCAGATGTTTCAAGTATGCTGCTGTGCTGTGGAAATTTCTCACT 1980  
Db 1921 CGGTACACCATCAAGCAGATGTTTCAAGTATGCTGCTGTGCTGTGGAAATTTCTCACT 1980  
QY 1981 GCGCAATTCATTCGCTCATCTCAGCCAGCGGCTCGGCGACAGACATGCTTACCA 2040  
Db 1981 GCGCAATTCATTCGCTCATCTCAGCCAGCGGCTCGGCGACAGACATGCTTACCA 2040

QY 2041 CACATCAGACTCCCATTTGGCTATTCAATCCCAAGCCCATATCATCTCTGCTGATAACA 2100  
Db 2041 CACATCAGACTCCCATTTGGCTATTCAATCCCAAGCCCATATCATCTCTGCTGATAACA 2100  
QY 2101 GGGTGGAAACGATCTCTGAAAGGAAGACCCGAATTTTCTGAAGTTGTCATGAAGTTAGAA 2160  
Db 2101 GGGTGGAAACGATCTCTGAAAGGAAGACCCGAATTTTCTGAAGTTGTCATGAAGTTAGAA 2160  
QY 2161 GAGTGTCTCTGACATTTGAGCTGATGTCCTGCAATCAAGTAACAGCAGTGGTCTCTC 2220  
Db 2161 GAGTGTCTCTGACATTTGAGCTGATGTCCTGCAATCAAGTAACAGCAGTGGTCTCTC 2220  
QY 2221 TCACCTTCTTCTTCTGATTTGCTGAAACCGGGGAGGACCTGGCCGAGTCAATG 2280  
Db 2221 TCACCTTCTTCTTCTGATTTGCTGAAACCGGGGAGGACCTGGCCGAGTCAATG 2280  
QY 2281 GCACATTAAGAGTCTGTTTTCGAATTTGGAATATGCTCTTAAATGCAAGTCCATGCTGCT 2340  
Db 2281 GCACATTAAGAGTCTGTTTTCGAATTTGGAATATGCTCTTAAATGCAAGTCCATGCTGCT 2340  
QY 2341 TTGTCCCAAAAGTGTGGACAATATTCCTCTCAAGTCTGTCTTTGGAGGAGATGAAAAGA 2400  
Db 2341 TTGTCCCAAAAGTGTGGACAATATTCCTCTCAAGTCTGTCTTTGGAGGAGATGAAAAGA 2400  
QY 2401 AGTCTTCAATACACACCATTTGAATAATGCTATGCTATCCGATCCCATGAGCTCAATG 2460  
Db 2401 AGTCTTCAATACACACCATTTGAATAATGCTATGCTATCCGATCCCATGAGCTCAATG 2460  
QY 2461 CATTTTCATTTCTGCGAAATAGTAGCAGCTTTGAGGACAGCAGC 2505  
Db 2461 CATTTTCATTTCTGCGAAATAGTAGCAGCTTTGAGGACAGCAGC 2505

## RESULT 4

US-09-291-839-1

; Sequence 1, Application US/09291839A

; Patent No. 6261818

; GENERAL INFORMATION:

; APPLICANT: Raju, Jeyaseelan

; TITLE OF INVENTION: NOVEL CARP PROTEIN AND NUCLEIC ACID MOLECULES AND USES THEREFOR

; FILE REFERENCE: WNI-068

; CURRENT APPLICATION NUMBER: US/09/291,839A

; CURRENT FILING DATE: 1999-04-14

; NUMBER OF SEQ ID NOS: 6

; SOFTWARE: PatentIn ver. 2.0

; SEQ ID NO 1

; LENGTH: 3025

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (48)..(2552)

US-09-291-839-1

Query Match

Best Local Similarity 100.0%; Score 2505; DB 3; Length 3025;

Matches 2505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGAATTTATAATCTAGACCAACCCAACTTGTACTGATGAATGGAAGAAAAAGTGC 60

Db 48 ATGGGAATTTATAATCTAGACCAACCCAACTTGTACTGATGAATGGAAGAAAAAGTGC 107

QY 61 AGTCAATCATCTTATCAATAGAAAGATTAGAGATGACCTGCAGATCAAGGAAAA 120

Db 108 AGTGAATCATATTTATCAATAGAAAGATTAGAGATGACCTGCAGATCAAGGAAAA 167

QY 121 GAAGTCAAGAACTAAGGAATATATTGGCTCTGTATGAAGCCCTTCAGTAAAGTCAATTTA 180

Db 168 GAAGTCAAGAACTAAGGAATATATTGGCTCTGTATGAAGCCCTTCAGTAAAGTCAATTTA 227

QY 181 AATTACCGCACTGAAAAATGGGCTGTCTTACTTCAATTATGTTGCAATTTGTGGAGGCAAG 240

Db 228 AATTACCGCACTGAAAAATGGGCTGTCTTACTTCAATTATGTTGCAATTTGTGGAGGCAAG 287

QY	241	AAATCACATATTCGAATCTCTTATGTTGAAAGGGCTCGCCCATCTCGACTGACAGAAAT	300	QY	1321	GAGAAGCAGATATTCCTCTCTAAGAGCTGGATTCCTTCAATTTCCATCTTCAGCTC	1380
Db	288	AAATCACATATTCGAATCTCTTATGTTGAAAGGGCTCGCCCATCTCGACTGACAGAAAT	347	Db	1368	GAGAAGCAGATATTCCTCTCTAAGAGCTGGATTCCTTCAATTTCCATCTTCAGCTC	1427
QY	301	GGATTACAGCCTTCATTTAGCAGTTTACAAGGATAATGACAGAAATGATCATCTCTCTG	360	QY	1381	TCAGAAATTCAGTTCCATGAGATTTATGGCTCAGGTTCTTTTGGGAAAGTATATAAGGA	1440
Db	348	GGATTACAGCCTTCATTTAGCAGTTTACAAGGATAATGACAGAAATGATCATCTCTCTG	407	Db	1428	TCAGAAATTCAGTTCCATGAGATTTATGGCTCAGGTTCTTTTGGGAAAGTATATAAGGA	1487
QY	361	CTTCACAGTGGAGTGATATACAGCAGGTTGGATACGGTGGCTCTCACTGGCCCTCCATAT	420	QY	1441	CGATGCGAAATAAATAGTGGCTATATAAAGCTTATCGAGCCCAATACCTACTGCTCCAA	1500
Db	408	CTTCACAGTGGAGTGATATACAGCAGGTTGGATACGGTGGCTCTCACTGGCCCTCCATAT	467	Db	1488	CGATGCGAAATAAATAGTGGCTATATAAAGCTTATCGAGCCCAATACCTACTGCTCCAA	1547
QY	421	GCTACAATAGCTGGCCACTAGAGGCTGCTGATGTGCTGTTGCAACATGAGAGCTAATGTC	480	QY	1501	TCAGATGATATGTTTTTCCGAGAGGTTCCATCTCTGCGAGCTCAATCATCCCTGC	1560
Db	468	GCTACAATAGCTGGCCACTAGAGGCTGCTGATGTGCTGTTGCAACATGAGAGCTAATGTC	527	Db	1548	TCAGATGATATGTTTTTCCGAGAGGTTCCATCTCTGCGAGCTCAATCATCCCTGC	1607
QY	481	AAATTTCAAGATGCAATTTTTCATCTCATATGATATGAGAGGTAATGAGACATGAA	540	QY	1561	GTAATTCAGTTTGGGCTGCTTGTGAATGATCCAGCCAGTTTTCCTACTCA	1620
Db	528	AAATTTCAAGATGCAATTTTTCATCTCATATGATATGAGAGGTAATGAGACATGAA	587	Db	1608	GTAATTCAGTTTGGGCTGCTTGTGAATGATCCAGCCAGTTTTCCTACTCA	1667
QY	541	CAGGTAACCTGCTCTTTTGAATTTGCTGCTGATGTAATGTAAGTGGTGAAGTTGGA	600	QY	1621	TACATATCAGGGGTTCTCTGTTCTCCCTCATGAGCAGAGGATTTCTGATTTG	1680
Db	588	CAGGTAACCTGCTCTTTTGAATTTGCTGCTGATGTAATGTAAGTGGTGAAGTTGGA	647	Db	1668	TACATATCAGGGGTTCTCTGTTCTCCCTCATGAGCAGAGGATTTCTGATTTG	1727
QY	601	GATAGACCCCTCCACTAGCATCTGCAAAAGGATCTTTGAATATGTCAAATCTTGATG	660	QY	1681	CAGTCTAAATTAATTTAGTAGATGTTGCCAAAGGATGGAGTACCTTCAACACTG	1740
Db	648	GATAGACCCCTCCACTAGCATCTGCAAAAGGATCTTTGAATATGTCAAATCTTGATG	707	Db	1728	CAGTCTAAATTAATTTAGTAGATGTTGCCAAAGGATGGAGTACCTTCAACACTG	1787
QY	661	GAGAGGCGACAGACAGATGTAATCTCAAGATATGAAGACCATGTCCTCACTCCAT	720	QY	1741	ACACAGCCAAATPACATCGTGAACAGTCAAAATTTCTTCTCTATGAGGATGG	1800
Db	708	GAGAGGCGACAGACAGATGTAATCTCAAGATATGAAGACCATGTCCTCACTCCAT	767	Db	1788	ACACAGCCAAATPACATCGTGAACAGTCAAAATTTCTTCTCTATGAGGATGG	1847
QY	721	TTCTGTCTCGATTTGGACACCATGATATAGTTAAGTATCTGTCGAAAAGTATTTGAA	780	QY	1801	CATCTGCTGGTGGCAGATTTTGGAGATCAAGATTTCTACAGTCTCTGGATGAAGCAAC	1860
Db	768	TTCTGTCTCGATTTGGACACCATGATATAGTTAAGTATCTGTCGAAAAGTATTTGAA	827	Db	1848	CATCTGCTGGTGGCAGATTTTGGAGATCAAGATTTCTACAGTCTCTGGATGAAGCAAC	1907
QY	781	GTTCACCTCATGTTGTTAATCTATGAGATACCCCTTACACCTGGCATCTACAAT	840	QY	1861	ATGCAAAACCAACCTGGGAACTCCGTTGGATGGCTCTGAGGTGTTACGAGTGCAT	1920
Db	828	GTTCACCTCATGTTGTTAATCTATGAGATACCCCTTACACCTGGCATCTACAAT	887	Db	1908	ATGCAAAACCAACCTGGGAACTCCGTTGGATGGCTCTGAGGTGTTACGAGTGCAT	1967
QY	841	GGCAATTTGAAGTTGCCAAGGAAATCATCCAAATATCAGGAAACAGAAAGTCTGCTAAG	900	QY	1921	CGGTACACCATCAAGCAGATGTTTACAGTATGCTCTGCTGCTGTTGGGAAATTTCTCACT	1980
Db	888	GGCAATTTGAAGTTGCCAAGGAAATCATCCAAATATCAGGAAACAGAAAGTCTGCTAAG	947	Db	1968	CGGTACACCATCAAGCAGATGTTTACAGTATGCTCTGCTGCTGTTGGGAAATTTCTCACT	2027
QY	901	GAAAACATCTTCAGTGAAACAGCTTTTCATAGTGCTTACCTATGCAAGAGCATTTGAC	960	QY	1981	GGCGAAATTCATTCGCTCATCTCAAGCCAGCGGCTGGGCGACAGACATGGCTTACCAC	2040
Db	948	GAAAACATCTTCAGTGAAACAGCTTTTCATAGTGCTTACCTATGCAAGAGCATTTGAC	1007	Db	2028	GGCGAAATTCATTCGCTCATCTCAAGCCAGCGGCTGGGCGACAGACATGGCTTACCAC	2087
QY	961	CTAGTCAAAATTTCTTCTGATCAGAAATGTCTAAACATCAACCAAGGAGGATGGG	1020	QY	2041	CACATCAGACCTCCATGGCTATTCATTCGAGCCCATATCTCTGCTGATACGA	2100
Db	1008	CTAGTCAAAATTTCTTCTGATCAGAAATGTCTAAACATCAACCAAGGAGGATGGG	1067	Db	2088	CACATCAGACCTCCATGGCTATTCATTCGAGCCCATATCTCTGCTGATACGA	2147
QY	1021	CACACTGGATTACACTCTGCTTGTCTACCAAGGATGATCCAGAGGCTAGTGGTGA	1080	QY	2101	GGGTGGACGATGCTCTGAGGAGGACCCGAAATTTCTGAAGTGTCTATGAAGTTAGAA	2160
Db	1068	CACACTGGATTACACTCTGCTTGTCTACCAAGGATGATCCAGAGGCTAGTGGTGA	1127	Db	2148	GGGTGGACGATGCTCTGAGGAGGACCCGAAATTTCTGAAGTGTCTATGAAGTTAGAA	2207
QY	1081	GATATGGAGCTGATATGATAGTGGCTTGTGATCCAGAGGCTAGTGGTGA	1140	QY	2161	GAGTGTCTCTGCAACATGAGTGTCTCTGCTGATCAAGTAAACAGCAGTGGGTCTCTC	2220
Db	1128	GATATGGAGCTGATATGATAGTGGCTTGTGATCCAGAGGCTAGTGGTGA	1187	Db	2208	GAGTGTCTCTGCAACATGAGTGTCTCTGCTGATCAAGTAAACAGCAGTGGGTCTCTC	2267
QY	1141	GATGACGACATGTTTGTATGGCTTATGAAAAGGCGATGATGCCATTTGACACTC	1200	QY	2221	TCACCTTCTCTTCTTCTGATGCTGTAACCGGGGAGCCTGCGGAGTCACTG	2280
Db	1188	GATGACGACATGTTTGTATGGCTTATGAAAAGGCGATGATGCCATTTGACACTC	1247	Db	2268	TCACCTTCTCTTCTTCTGATGCTGTAACCGGGGAGCCTGCGGAGTCACTG	2327
QY	1201	CTGAGCATTAAGAGACCAAGATGATGCTGTAATGCTGTAATGATATTTCTCAGCTTGA	1260	QY	2281	CGACATTAAGAGAGTGGTTTCGAATTTGGAATATGCTTAAATGCAAGTCTATGCTGCT	2340
Db	1248	CTGAGCATTAAGAGACCAAGATGATGCTGTAATGCTGTAATGATATTTCTCAGCTTGA	1307	Db	2328	CGACATTAAGAGAGTGGTTTCGAATTTGGAATATGCTTAAATGCAAGTCTATGCTGCT	2387
QY	1261	GGAGATGGCTCTCTATGCTGTTCCATCACCTTTGGGAAAGATTAAGAGCATGACAAA	1320	QY	2341	TTGTCCCAAGTGTGACATATTTCTCTCAGAGTCTGCTTTGGAGGAGATGAAGA	2400
Db	1308	GGAGATGGCTCTCTATGCTGTTCCATCACCTTTGGGAAAGATTAAGAGCATGACAAA	1367	Db	2388	TTGTCCCAAGTGTGACATATTTCTCTCAGAGTCTGCTTTGGAGGAGATGAAGA	2447
				QY	2401	AGTCTTCAATACACACCCATTTGACAAATATGGCTATGATCCGATCCCATGCTCAATG	2460

Db 2448 AGTCTTCATACACCCATGACAAATATGGCTATGTATCCCATGCCATGAGCTCAATG 2507  
Qy 2461 CATTTTCATCTTCCGAAATAGTAGCAGCTTTGAGCAGCAGC 2505  
Db 2508 CATTTTCATCTTCCGAAATAGTAGCAGCTTTGAGCAGCAGC 2552

RESULT 5  
US-09-458-457-1  
; Sequence 1, Application US/09458457  
; Patent No. 6500654  
; GENERAL INFORMATION:  
; APPLICANT: Raju, Jeyaseelan  
; TITLE OF INVENTION: NOVEL CARP PROTEIN AND NUCLEIC ACID MOLECULES AND USES  
; TITLE OF INVENTION: THEREFOR  
; FILE REFERENCE: mni-068cp  
; CURRENT APPLICATION NUMBER: US/09/458,457  
; CURRENT FILING DATE: 1999-12-10  
; EARLIER APPLICATION NUMBER: 60/111,938  
; EARLIER FILING DATE: 1998-12-11  
; EARLIER APPLICATION NUMBER: 09/291,839  
; EARLIER FILING DATE: 1999-04-14  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 3025  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (48)..(2552)  
US-09-458-457-1

Query Match 100.0%; Score 2505; DB 4; Length 3025;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGGGAAATTAATACTAGACCAACCCAACTTGTAATGATGGAAGGAAAGGTC 60  
Db 48 ATGGGAAATTAATACTAGACCAACCCAACTTGTAATGATGGAAGGAAAGGTC 107

Qy 61 AGTGAATCATATGTTATCAACATAGAAAGATGAGAGATGACCTGAGATCAAGGAAAA 120  
Db 108 AGTGAATCATATGTTATCAACATAGAAAGATGAGAGATGACCTGAGATCAAGGAAAA 167

Qy 121 GAACGTGACAGAACTAAGGAATATTTGGCTCTGATGAAGCCCTTCAGTAAAGTCAATTTA 180  
Db 168 GAACGTGACAGAACTAAGGAATATTTGGCTCTGATGAAGCCCTTCAGTAAAGTCAATTTA 227

Qy 181 AATTACCGCACTGAAATGGGCTGCTCTACTTCATTATGTTGCAATTTGTCGAGGCAAG 240  
Db 228 AATTACCGCACTGAAATGGGCTGCTCTACTTCATTATGTTGCAATTTGTCGAGGCAAG 287

Qy 241 AAATCACATATTCGAATCTTATGTTGAAGGGCTCCGCCATCTCGACTGACAGAAAT 300  
Db 288 AAATCACATATTCGAATCTTATGTTGAAGGGCTCCGCCATCTCGACTGACAGAAAT 347

Qy 301 GGATTTACAGCCTTGCATTTAGCAGTTTACAAGGATTAATGCAAGATGATCACTTCTCTG 360  
Db 348 GGATTTACAGCCTTGCATTTAGCAGTTTACAAGGATTAATGCAAGATGATCACTTCTCTG 407

Qy 361 CTTTCAGAGTGGAGTGNATATACAGAGGTTGATAGGCTGAGGCTGAGGCTCCATATT 420  
Db 408 CTTTCAGAGTGGAGTGNATATACAGAGGTTGATAGGCTGAGGCTGAGGCTCCATATT 467

Qy 421 GCTACAAATAGTGGCACTAGAGGCTGCTGATGCTGTGCAATGCAATGAGCTAATGTC 480  
Db 468 GCTACAAATAGTGGCACTAGAGGCTGCTGATGCTGTGCAATGCAATGAGCTAATGTC 527

Qy 481 AATATTCAGATGAGTTTTCATCCATGCAATTTGCAAGGCTATGAGCATGAA 540  
Db 528 AATATTCAGATGAGTTTTCATCCATGCAATTTGCAAGGCTATGAGCATGAA 587

Qy 541 CAGGTAACTCGCCTTCTTTGAAATTTGGTGCTGATGTAAATGTAAAGTGGTGAAGTTGA 600  
Db 588 CAGGTAACTCGCCTTCTTTGAAATTTGGTGCTGATGTAAATGTAAAGTGGTGAAGTTGA 647  
Qy 601 GATAGACCCCTCCACCTAGCATCTGCAAAAGGATTTCTGAATATTGCAAAATCTTTGATG 660  
Db 648 GATAGACCCCTCCACCTAGCATCTGCAAAAGGATTTCTGAATATTGCAAAATCTTTGATG 707  
Qy 661 GAAGAAGGAGCAAAAGCAGATGTGAATGCTCAAGATAATGAAGACCATGCTCCACTCCAT 720  
Db 708 GAAGAAGGAGCAAAAGCAGATGTGAATGCTCAAGATAATGAAGACCATGCTCCACTCCAT 767  
Qy 721 TTCTGTCTCGATTTGCAACACATGATATAGTTAAAGTATCTGCTGCAAAAGTGAATGGA 780  
Db 768 TTCTGTCTCGATTTGCAACACATGATATAGTTAAAGTATCTGCTGCAAAAGTGAATGGA 827  
Qy 781 GTTCAACCTCATGTTGTTAATATCTATGGAGATACCCCTTACACTGCGCATGCTACCAT 840  
Db 828 GTTCAACCTCATGTTGTTAATATCTATGGAGATACCCCTTACACTGCGCATGCTACCAT 887  
Qy 841 GGCATAATTTGAAGTGGCAAGGAAATCATCCAAATATCAGGAACAGAAAGTCTGACTAAG 900  
Db 888 GGCATAATTTGAAGTGGCAAGGAAATCATCCAAATATCAGGAACAGAAAGTCTGACTAAG 947  
Qy 901 GAAAAACATCTTCAGTGAACACAGCTTTTCATAGTGTGTACCTATGCGAAGACATTCAC 960  
Db 948 GAAAAACATCTTCAGTGAACACAGCTTTTCATAGTGTGTACCTATGCGAAGACATTCAC 1007  
Qy 961 CTAGTCAAAATTTCTTTGATCAGAAATGTCAATAACATCAACCAAGGAAGGATGGG 1020  
Db 1008 CTAGTCAAAATTTCTTTGATCAGAAATGTCAATAACATCAACCAAGGAAGGATGGG 1067  
Qy 1021 CACACTGGATTAACACTCTGCTTACACAGCTGCATTCGCTCGCTTCACTTCTTACTG 1080  
Db 1068 CACACTGGATTAACACTCTGCTTACACAGCTGCATTCGCTCGCTTCACTTCTTACTG 1127  
Qy 1081 GATAATGGAGCTGATATGAATCTAGTGGCTTGTGATCCAGCAGGTCTAGTGGTGAAGAA 1140  
Db 1128 GATAATGGAGCTGATATGAATCTAGTGGCTTGTGATCCAGCAGGTCTAGTGGTGAAGAA 1187  
Qy 1141 GATGAGCAGACATGTTGATGAGGCTTATGAAAAGGCGATGATGCCATTTGTCACTC 1200  
Db 1188 GATGAGCAGACATGTTGATGAGGCTTATGAAAAGGCGATGATGCCATTTGTCACTC 1247  
Qy 1201 CTGAAGCATTTAAGAGACCAAGATGAATTCCTGCTGTAATGAATATTTCTCAGCTCGA 1260  
Db 1248 CTGAAGCATTTAAGAGACCAAGATGAATTCCTGCTGTAATGAATATTTCTCAGCTCGA 1307  
Qy 1261 GGAGATGGCTCTTATGCTGTTCCATCACCTTGGGGAAGATTAAGCATGACAAAA 1320  
Db 1308 GGAGATGGCTCTTATGCTGTTCCATCACCTTGGGGAAGATTAAGCATGACAAAA 1367  
Qy 1321 GAGAAGGAGATATTCCTCTTAAGAGCTGGATTCCTTCACTTTCATCTTCCAGCTC 1380  
Db 1368 GAGAAGGAGATATTCCTCTTAAGAGCTGGATTCCTTCACTTTCATCTTCCAGCTC 1427  
Qy 1381 TCAGAAATTTGAGTTCATGAGATTTATGGCTCAGGTTCTTTTGGGAAAGTATATAAGGA 1440  
Db 1428 TCAGAAATTTGAGTTCATGAGATTTATGGCTCAGGTTCTTTTGGGAAAGTATATAAGGA 1487  
Qy 1441 CGATGAGAAATAAATAGTGGCTATAAAGCTTATCGAGCCAAATACCTACTGCTCGAAG 1500  
Db 1488 CGATGAGAAATAAATAGTGGCTATAAAGCTTATCGAGCCAAATACCTACTGCTCGAAG 1547  
Qy 1501 TCAGATGTGGATATGTTTGGCGAGAGGTTGCTTCTCTGCGAGCTCAATCATCCCTGC 1560  
Db 1548 TCAGATGTGGATATGTTTGGCGAGAGGTTGCTTCTCTGCGAGCTCAATCATCCCTGC 1607  
Qy 1561 GTAATTCAGTTTGTGGTCTCTGCTTGAATGATCCAGCCAGTTTGGCATTTGTCTACTCA 1620  
Db 1608 GTAATTCAGTTTGTGGTCTCTGCTTGAATGATCCAGCCAGTTTGGCATTTGTCTACTCA 1667



Db 828 GTTCAACTCATGTTGTTAATATCTATGAGATACCCCTTACACCTGCATGCTACAAT 887  
Qy 841 GGCAGAAATTTGAAGTGCAGAGGAAATCATCAATATCAGGACAGAAAGTCTGACTAAG 900  
Db 888 GGCAGAAATTTGAAGTGCAGAGGAAATCATCAATATCAGGACAGAAAGTCTGACTAAG 947  
Qy 901 GAAACATCTTCAGTGAAGACAGCTTTTCATAGTGTGTACCTATGCGCAAGAGCATTCAC 960  
Db 948 GAAACATCTTCAGTGAAGACAGCTTTTCATAGTGTGTACCTATGCGCAAGAGCATTCAC 1007  
Qy 961 CTAGTCAAAATCTTCCTGATCAGATGTCAATAAACATCAACACCAAGAGGAGGATGGG 1020  
Db 1008 CTAGTCAAAATCTTCCTGATCAGATGTCAATAAACATCAACACCAAGAGGAGGATGGG 1067  
Qy 1021 CACACTGATTTACACTCTGCTTCCACCGGTACATTCGCTGCTTTCAGTCTTCTACTG 1080  
Db 1068 CACACTGATTTACACTCTGCTTCCACCGGTACATTCGCTGCTTTCAGTCTTCTACTG 1127  
Qy 1081 GATAATGAGCTGATATGAATAGTGGCTTGTGATCCCAGCAGGTCTAGTGGTGAAGAA 1140  
Db 1128 GATAATGAGCTGATATGAATAGTGGCTTGTGATCCCAGCAGGTCTAGTGGTGAAGAA 1187  
Qy 1141 GATGAGCAGACATGTTTGATGTGGCTTATGAAGAGGACATGATCCCATTCACACTC 1200  
Db 1188 GATGAGCAGACATGTTTGATGTGGCTTATGAAGAGGACATGATCCCATTCACACTC 1247  
Qy 1201 CTGAAGCATTTAAGAGACACCAAGATGAATTCCTGTAATGAATATTCCTCAGCTCGA 1260  
Db 1248 CTGAAGCATTTAAGAGACACCAAGATGAATTCCTGTAATGAATATTCCTCAGCTCGA 1307  
Qy 1261 GGAGATGGCTCTATGCTGTCTGATCCCTTGGGAGAGATTAAGAGCATGACAAA 1320  
Db 1308 GGAGATGGCTCTATGCTGTCTGATCCCTTGGGAGAGATTAAGAGCATGACAAA 1367  
Qy 1321 GAGAGGCAGATATTCCTCTTAAGAGCTGATGCTTCAATTTCCATCTCAGCTC 1380  
Db 1368 GAGAGGCAGATATTCCTCTTAAGAGCTGATGCTTCAATTTCCATCTCAGCTC 1427  
Qy 1381 TCAGAAATGAGTTCATGAGATATTTGGCTCAGTTCATTTTGGGAAAGTATATAAGGA 1440  
Db 1428 TCAGAAATGAGTTCATGAGATATTTGGCTCAGTTCATTTTGGGAAAGTATATAAGGA 1487  
Qy 1441 CGATGCGAATATAATAGTGGCTTATAAGAGTATCGGCAATACCTACTGCTCCAG 1500  
Db 1488 CGATGCGAATATAATAGTGGCTTATAAGAGTATCGGCAATACCTACTGCTCCAG 1547  
Qy 1501 TCAGATGTGATATGTTTCCGAGAGGTGTCCATTCCTGCGAGCTCAATCATCCCTGC 1560  
Db 1548 TCAGATGTGATATGTTTCCGAGAGGTGTCCATTCCTGCGAGCTCAATCATCCCTGC 1607  
Qy 1561 GTAAATCAGTTTGGGTGCTTGTGATGATCCAGCCAGTTGCCATTTGCTACTCA 1620  
Db 1608 GTAAATCAGTTTGGGTGCTTGTGATGATCCAGCCAGTTGCCATTTGCTACTCA 1667  
Qy 1621 TACATATCAGGGGTCTCTGTTCTCCCTCTTCATGAGCAGAGAGGATTTCTGATTTG 1680  
Db 1668 TACATATCAGGGGTCTCTGTTCTCCCTCTTCATGAGCAGAGAGGATTTCTGATTTG 1727  
Qy 1681 CAGTCTAAATTAATATTGAGTAGATGTTGCCAAAGGATGAGTACCTTCAACACCTG 1740  
Db 1728 CAGTCTAAATTAATATTGAGTAGATGTTGCCAAAGGATGAGTACCTTCAACACCTG 1787  
Qy 1741 ACAGAGCAATATACATCGTGAATGAGCAGTCAATATTCCTCTATGAGGATGG 1800  
Db 1788 ACAGAGCAATATACATCGTGAATGAGCAGTCAATATTCCTCTATGAGGATGG 1847  
Qy 1801 CATGCTGTGTGGCAGATTTTGGAGATCAAGATTTCTCAGTCTCTGGATGAAGACAC 1860  
Db 1848 CATGCTGTGTGGCAGATTTTGGAGATCAAGATTTCTCAGTCTCTGGATGAAGACAC 1907  
Qy 1861 ATGACAAAACAACTGGACCTCCGTTGGATGGCTCTGAGGTGTTACGAGTGCAT 1920

Db 1908 ATGACAAAACAACTGGGAACTCGTTGGATGGCTCTCGAGTGTTCACGAGTGCAT 1967  
Qy 1921 CGGTACACCATCAAGCAGATGCTTTCAGCTATGCTCTGTGTCTGTGGAAATTCCTACT 1980  
Db 1968 CGGTACACCATCAAGCAGATGCTTTCAGCTATGCTCTGTGTCTGTGGAAATTCCTACT 2027  
Qy 1981 GGGGAAATTCATTCCTCATCTCAAGCCAGCGGTGCGGACAGACATGCTTACCAC 2040  
Db 2028 GGGGAAATTCATTCCTCATCTCAAGCCAGCGGTGCGGACAGACATGCTTACCAC 2087  
Qy 2041 CACATCAGACCTCCCATTTGGCTATTCATTCACAGCCCATATCATCTCTGCTGATAGA 2100  
Db 2088 CACATCAGACCTCCCATTTGGCTATTCATTCACAGCCCATATCATCTCTGCTGATAGA 2147  
Qy 2101 GGGTGGAAACGATGCTCCTGAAAGAGACCCGAAATTTCTGAAGTGTCTCATGAAGTTAGAA 2160  
Db 2148 GGGTGGAAACGATGCTCCTGAAAGAGACCCGAAATTTCTGAAGTGTCTCATGAAGTTAGAA 2207  
Qy 2161 GAGTGTCTCTGAAACATTCAGCTGATGCTCTCTGATCAAGTAAACAGAGTGGTCTCTC 2220  
Db 2208 GAGTGTCTCTGAAACATTCAGCTGATGCTCTCTGATCAAGTAAACAGAGTGGTCTCTC 2267  
Qy 2221 TCACCTTCTTCTTCTTCTGATTCCTGATTCCTGATTCCTGATTCCTGATTCCTGATTCCTG 2280  
Db 2268 TCACCTTCTTCTTCTTCTGATTCCTGATTCCTGATTCCTGATTCCTGATTCCTGATTCCTG 2327  
Qy 2281 GCAGCATTAAGAACTGCTTTCGAATTCGAATTCGAATTCGAATTCGAATTCGAATTCGAATTCGA 2340  
Db 2328 GCAGCATTAAGAACTGCTTTCGAATTCGAATTCGAATTCGAATTCGAATTCGAATTCGAATTCGA 2387  
Qy 2341 TTGTCCCAAGTCTCGCAATATTCCTCTCAAGTCTCTCTTTCGAGGAGATGAAAGA 2400  
Db 2388 TTGTCCCAAGTCTCGCAATATTCCTCTCAAGTCTCTCTTTCGAGGAGATGAAAGA 2447  
Qy 2401 AGTCTTCAATACACACCCATTCGAATTCGAATTCGAATTCGAATTCGAATTCGAATTCGAATTCGA 2460  
Db 2448 AGTCTTCAATACACACCCATTCGAATTCGAATTCGAATTCGAATTCGAATTCGAATTCGAATTCGA 2507  
Qy 2461 CATTTTCATTCCTGCGCAATATGATAGCAGCTTTGAGGACAGCAGC 2505  
Db 2508 CATTTTCATTCCTGCGCAATATGATAGCAGCTTTGAGGACAGCAGC 2552

RESULT 7

US-09-458-457-9  
; Sequence 9, Application US/09458457  
; Patent No. 6500654  
; GENERAL INFORMATION:  
; APPLICANT: Raju, Jeyaseelan  
; TITLE OF INVENTION: NOVEL CARP PROTEIN AND NUCLEIC ACID MOLECULES AND USES  
; TITLE OF INVENTION: THEREFOR  
; FILE REFERENCE: mni-068cp  
; CURRENT APPLICATION NUMBER: US/09/458,457  
; CURRENT FILING DATE: 1999-12-10  
; EARLIER APPLICATION NUMBER: 60/111,938  
; EARLIER FILING DATE: 1998-12-11  
; EARLIER APPLICATION NUMBER: 09/291,839  
; EARLIER FILING DATE: 1999-04-14  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: PatentIn ver. 2.0  
; SEQ ID NO 9  
; LENGTH: 2505  
; TYPE: DNA  
; ORGANISM: Rattus norvegicus  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(2505)  
US-09-458-457-9

Query Match 75.5%; Score 1892.2; DB 4; Length 2505;  
Best Local Similarity 84.7%; Pred. No. 0;  
Matches 2122; Conservative 0; Mismatches 383; Indels 0; Gaps 0;



Db 2161 GAGTGCCTATCAATGTGGAGCTCATGTCTCCAGCATCAAGTAACAGCAGTGGCTCTCTG 2220  
Qy 2221 TCACCTTCTTCTTCTCTGTGATGCTGTGTAACCGGGGAGACCTGGCGGAGTCATGTG 2280  
Db 2221 TCACCTTCTTCTTCTTCCGATGCTGTGAGCGGGGAGGCGCTGGCGGAGCAACGTG 2280  
Qy 2281 GCAGCATTAAGAAAGTGGTTCGAAATGGAAATGATGCTCTAAATGCAAGGTCTCTATGCTCT 2340  
Db 2281 GCAGCCTTACGGAGCGGTTTGGATTGGAGTATGCCCTAAATGCAAGGTCTCTATGCTGG 2340  
Qy 2341 TTGTCCCAAGTCTGGACAATATCTCTCAAGGTCTCTCTTTGGAGAGAGTAAAGA 2400  
Db 2341 TGGTCCCAAGTCTGGACAATATCTCTCAAGGTCTCTCTTTGGAGAGATGAATAGG 2400  
Qy 2401 AGTCTCAATACACACCATGACCAATATGCTATGATCGATCCGATGAGTCAATG 2460  
Db 2401 AGCACCAGTATTCAACTGTGCAATACGGGTATGCTGATCCCATGAGCCTGAGC 2460  
Qy 2461 CATTTTCATTTGCCGAATATGACAGCTTTGAGGACAGCAGC 2505  
Db 2461 CACCTTCACTCCCGCCAAGACGACGCAACTTTGAGACAGCAAC 2505

RESULT 8  
US-09-947-199A-9  
; Sequence 9, Application US/09947199A  
; Patent No. 6660490  
; GENERAL INFORMATION:  
; APPLICANT: Raju, Jeyaseelan  
; TITLE OF INVENTION: NOVEL CARP PROTEIN AND NUCLEIC ACID MOLECULES AND USES  
; FILE REFERENCE: THEREFOR  
; CURRENT APPLICATION NUMBER: US/09/947,199A  
; CURRENT FILING DATE: 2001-09-05  
; PRIOR FILING DATE: US 60/111,938  
; PRIOR FILING DATE: 1998-12-11  
; PRIOR APPLICATION NUMBER: US 09/291,839  
; PRIOR FILING DATE: 1999-04-14  
; PRIOR APPLICATION NUMBER: US 09/458,457  
; PRIOR FILING DATE: 1999-12-10  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 9  
; LENGTH: 2505  
; TYPE: DNA  
; ORGANISM: Rattus norvegicus  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(2505)  
US-09-947-199A-9

Query Match 75.5%; Score 1892.2; DB 4; Length 2505;  
Best Local Similarity 84.7%; Pred. No. 0;  
Matches 2122; Conservative 0; Mismatches 383; Indels 0; Gaps 0;

Qy 1 ATGGGAATTAATAAATCTAGACCAACCCAACTTCTACTGATGAATGGAAGAAAATGTC 60  
Db 1 ATGGGAATTAATAAATCTAGACCAACACACAGACTTGTCTGATGAATGGAAGAAAATGTC 60  
Qy 61 AGTGAATCATATGTTATCAAAATAGAAAGATTAGAGATGACCTGCAGATCAAGGAAAA 120  
Db 61 AGTGAATCATACCTATTATCATAAGAGGCTGGAGGATAACCTGCAGATCAAGAAAAAT 120  
Qy 121 GAATGACAGAACTAGGAATATATTTGGCTCTGATGAAGCCCTTCAGTAAAGTCAATTTA 180  
Db 121 GAATTTCAAGAACTAGAGGACATCTTTGGCTCTGATGAAGCCCTTCAGTAAAGTCAATTTA 180  
Qy 181 AATTACCGCACTGAAAATGGGCTGCTCTACTTTCATTTATGTTGCAATTTGTGGAGGCAAG 240  
Db 181 AATTACCGCACTAGAGGCTGCTCTACTTTCATTTATGTTGCAATTTGTGGAGGCAAG 240  
Qy 241 AATATCATATTCGAATCTTATGTTGAAAGGGCTCCGCGCAATCTGACTGACAAAGAAAT 300  
Db 241 AATATCATATTCGAATCTTATGTTGAAAGGGCTCCGCGCAATCTGACTGACAAAGAAAT 300

Db 241 AAGTCACATATCGTGCCTTATGTTAAAGGGCTCCGTCATCCAGACTGACGAGAAAT 300  
Qy 301 GGATTTACAGCCTTGCAATTTACAGTTTACAGAGTAATGACAGAAATGATCATTCTCTG 360  
Db 301 GGGTTTCCAGCTCTGCACCTGGCCGTTTACAGAGGACAGCCCGGAACTTATCATTCACTG 360  
Qy 361 CTTTCAGATGGAGCTGATATACAGCAGGTTGGATACGGTGGCTCACTGCCCTCCATATT 420  
Db 361 TTGCA CAGCGGAGCAGATGTTT CAGCAAGTGGATACGGTGGCTT CACAGCCCTCCACATA 420  
Qy 421 GCTACAAATAGCTGGCACTAGAGCTGCTGATGCTGTTGCAACATGGAGCTTAATGTC 480  
Db 421 GCTGCAATAGCTGGACACCCAGAGCTGCAGAAAGTGTCTGCTACAAACATGGGCGCAATGTG 480  
Qy 481 AATATTCAAGATGCGAGTTTTTTTTCATCCATTCATGATATTCAGAGCTACTATGGCAGCG 540  
Db 481 AATGTTCAAGATGCGGCTCTTCTTACCACCTGCACATTCAGAGCTACTATGGCAGCG 540  
Qy 541 CAGGTAACCTCGCCTTCTTTTGAATTTGGTGTGATGTAATGTAAAGTGGTGAAGTTGGA 600  
Db 541 CAGGTAACCTCGCCTTCTTTTGAATTTGGTGTGATGTAATGTAAAGCGGTGAAGTTGGG 600  
Qy 601 GATAGACCCCTCCACCTAGCATCTGCAAAAGGATTTTGAATATTGCAAACTCTTGATG 660  
Db 601 GACAGCCCTCTGCACCTGGCCCTTGCAGAGGGCTTCTTCAACATTTGAAAACTCCTGGTA 660  
Qy 661 GAAGAAGGACGCAAAAGCAGATGTAATGCTCAAGATAATGAAGACCATGTCCCACTCCAT 720  
Db 661 GAAGAAGGACGCAAAAGCAGATGTAATGCTCAAGATAATGAAGACCATGTCCCACTCCAT 720  
Qy 721 TTCTGTCTCGATTTGGACACCATGATATAGTTAAGTATCTGCTGCAAGTGAATTTGAA 780  
Db 721 TTCTGTCTCGATTTGGACACCATGATATAGTTAAGTATCTGCTGCAAGTGAATTTGAA 780  
Qy 781 GTTCAACCTCATGTTGTTAATATCTATGGAGATACCCCTTACACCTGGCATGTACAAAT 840  
Db 781 GTTCAACCTCATGTTGTTAATATCTATGGAGATACCCCTTACACCTGGCATGTACAAAT 840  
Qy 841 GGCATAATTTGAAGTTGCCAAGGAATATCCAAATATCAGGAACAAGAAAGTCTGACTAAG 900  
Db 841 GGCATAATTTGAAGTTGCCAAGGAATATCCAAAGTAAACAGGAACTGAAAGTCTGACTAAG 900  
Qy 901 GAAACATCTTCAGTGAACAGCTTTTCATAGTGTGCTGCTACTATGCGCAAGAGATTCAC 960  
Db 901 GAAACATCTTCAGTGAACAGCTTTTCATAGTGTGCTGCTACTATGCGCAAGAGATTCAC 960  
Qy 961 CTAGTCAAAATTTCTTCTTGATCAGAAATGTCATAAATCAACCAAGGAGGAGATGGG 1020  
Db 961 CTAGTCAAAATTTCTTCTTGATCAGAAATGTCATAAATCAACCAAGGAGGAGATGGG 1020  
Qy 1021 CACACTGGATACACTCTGCTTGTACACCGTACATTCGCTGGTTCAGTTCTTACTG 1080  
Db 1021 CACACTGGATACACTCTGCTTGTACACCGGCTATCCGCGCCATATCCGCTGGTTCAGTTCTTACTT 1080  
Qy 1081 GATAATGGAGCTGATATGAATCTAGTGGCTTATGAAAGGGCATGATGCCATTTGTCAGCTG 1140  
Db 1081 GATAATGGAGCTGATATGAATCTAGTGGCTTATGAAAGGGCATGATGCCATTTGTCAGCTG 1140  
Qy 1141 GATGACAGACATGTTGATGTTGAGTGGCTTATGAAAGGGCATGATGCCATTTGTCAGCTG 1200  
Db 1141 GATGACAGACATGTTGATGTTGAGTGGCTTATGAAAGGGCATGATGCCATTTGTCAGCTG 1200  
Qy 1201 CTGAGCATTTAAGAGACCAAGATGAATTCGCTGTAATGAATATTCCTCAGCTGGA 1260  
Db 1201 CTGAGCATTTAAGAGACCAAGATGAATTCGCTGTAATGAATATTCCTCAGCTGGA 1260  
Qy 1261 GGAGATGGCTCTCTATGTTCTTCCATCACCCTTTGGGAAAGATTTAAAGCATGACAAAA 1320  
Db 1261 GGAGATGGCTCTCTATGTTCTTCTTCCCTTTGGGCAAGATTTAAAGCATGACAAAA 1320  
Qy 1321 GAGAGGCGAGATTTCTCTCTCTAAGAGCTGATGCTTCAATTTCCATTTCCATTTCAAGCTC 1380  
Db 1321 GAGAGGCGAGATTTCTCTCTCTAAGAGCTGATGCTTCAATTTCCATTTCCATTTCAAGCTC 1380

```
QY 1381 TCAGAAATTGAGTTCATGATGATTTGGCTCAGGTTCTTTTGGGAAAGTATATAAGA 1440
Db 1381 TCCGAAATCGAGTTCACGAGATTTATCGGCTCGGTTCTTTGGGAAAGTCTATARAAGG 1440
QY 1441 CGATCAGAAATAAATAGTGGCTATAAAAGTTATCGAGCCAAATACCTACTGCTCCAAG 1500
Db 1441 CGATCAGAAATAAATAGTGGCTATAAAAGTTATCGGCTCGGTTCTTTGGGAAAGTCTATARAAGG 1500
QY 1501 TCAGATGAGATATGTTTCCGAGAGGTGTCCATCTCTGCCAGTCAATCATCCCTCG 1560
Db 1501 TCAGACGTGGATATGTTTCCGAGAGGTGTCCATCTCTGCCAGTCAATCATCCCTCG 1560
QY 1561 GTAATTCAGTTTGTGGTGCTTGTGTAATGATCCAGCCAGTTTCCCAATGTCACAA 1620
Db 1561 GTGGTTCAAGTTTGTGGTGCTGCTGGATGATCCAGTCAAGTTTCCCAATGTCACAA 1620
QY 1621 TACATATCAGGGGTTCTGTTCTCCCTCTCATGAGCAGAGGATTTCTTGATTG 1680
Db 1621 TACATATCAGGGGTTCTGTTCTCCCTCTCATGAGCAGAGGATTTCTTGATTG 1680
QY 1681 CAGTCTAAATTAATTCAGTATGATGTTGCCAAAGGCATGGAGTACCTTCACAACTG 1740
Db 1681 CAGTCTAAATTAATTCAGTATGATGTTGCCAAAGGCATGGAGTACCTTCACAACTG 1740
QY 1741 ACACAGCAATATACATCGTACTTGAACAGTCAATATTTCTCTATGAGGATGG 1800
Db 1741 ACCAGCAATATACATCGTACTTGAACAGTCAATATTTCTCTATGAGGATGG 1800
QY 1801 CATGCTGTGGGAGATTTGGAGATCAAGATTTCTACAGTCTCTGGATGAAGACAAC 1860
Db 1801 CATGCTGTGGGAGATTTGGAGATCAAGATTTCTACAGTCTCTGGATGAAGACAAC 1860
QY 1861 ATGACAAACCACTGGGAACCTCCGTTGGATGGCTCCTCAGTGTTCACAGTGCAT 1920
Db 1861 ATGACAAACCACTGGGAACCTCCGTTGGATGGCTCCTCAGTGTTCACAGTGCAT 1920
QY 1921 CGGTACACATCAAGCAGATGCTTCAGTATGCTGTCTGTCTGTGGGAAATTCATCT 1980
Db 1921 AGATACACATCAAGCAGATGCTTCAGTATGCTGTCTGTCTGTGGGAGTCTCTCAT 1980
QY 1981 GCGAAATTCATTCGCTCATCTCAAGCCAGCGGCTCGGAGCAGACATGGTTTACCAC 2040
Db 1981 GCGAAATTCATTCGCTCATCTCAAGCCAGCGGCTCGGAGCAGACATGGGTTATCAC 2040
QY 2041 CACATCAGACCTCCATGGCTATTCATTCAGGAGGAGGAGGAGGAGGAGGAGGAGG 2100
Db 2041 CACATCAGACCTCCATGGCTATTCATTCAGGAGGAGGAGGAGGAGGAGGAGGAGG 2100
QY 2101 GGGTGAACGATGCTCTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2160
Db 2101 GGGTGAACGATGCTCTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2160
QY 2161 GAGTGTCTGCAACATGAGTGTCTCTGATCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 2220
Db 2161 GAGTGTCTGCAACATGAGTGTCTCTGATCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 2220
QY 2221 TCACCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2280
Db 2221 TCACCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2280
QY 2281 GCAGCATTAAGAGTGTGTTTGAATATGCTCTTAAATCAAGGTCCTATGCTGT 2340
Db 2281 GCAGCATTAAGAGTGTGTTTGAATATGCTCTTAAATCAAGGTCCTATGCTGT 2340
QY 2341 TTGTCCTCAAGTGTGAGCAATATTCCTCAAGGTCCTCTTGGAGGAGTGAAGA 2400
Db 2341 TTGTCCTCAAGTGTGAGCAATATTCCTCAAGGTCCTCTTGGAGGAGTGAAGA 2400
QY 2401 AGTCTTCAATACACACCAATATGATGCTATGATCCGATCCCATGAGTCAATG 2460
Db 2401 AGTCTTCAATACACACCAATATGATGCTATGATCCGATCCCATGAGTCAATG 2460
```

```
QY 2461 CATTTTCATTCTTCCGAAATAGTAGCAGCTTTGAGGACAGCAGC 2505
Db 2461 CACCTTCACTCCCGCAAGACGACGACCAACTTTGAGGACAGCAAC 2505
```

## RESULT 9

```
US-09-458-457-7
; Sequence 7, Application US/09458457
; Patent No. 650654
; GENERAL INFORMATION:
; APPLICANT: Raju, Jeyaseelan
; TITLE OF INVENTION: NOVEL CARK PROTEIN AND NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: mni-068cp
; CURRENT APPLICATION NUMBER: US/09/458,457
; EARLIER FILING DATE: 1999-12-10
; EARLIER APPLICATION NUMBER: 60/111,938
; EARLIER FILING DATE: 1998-12-11
; EARLIER APPLICATION NUMBER: 09/291,839
; EARLIER FILING DATE: 1999-04-14
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 3026
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (61)..(2565)
US-09-458-457-7
```

```
Query Match 75.5%; Score 1892.2; DB 4; Length 3026;
Best Local Similarity 84.7%; Pred. No. 0;
Matches 2122; Conservative 0; Mismatches 383; Indels 0; Gaps 0;
QY 1 ATGGGAAATTAATACTAGACCAACCCAACTTGTACTGATGAATGAAGAAAAATGC 60
Db 61 ATGGGAAATTAATACTAGACCAACCCAACTTGTACTGATGAATGAAGAAAAATGC 120
QY 61 AGTGAATCATATGTTATCAATAGAAAGATAGAGATGACCTCAGATCAAGAAAAA 120
Db 121 AGTGAATCATATGTTATCAATAGAAAGCTGGAGATTAACCTCGAGATCAAGAAAAA 180
QY 121 GAACTGACAGAACTAAGGAATATATTGGCTCTGATGAAGCCTTCAGTAAAGTCAATT 180
Db 181 GAAATTCAGAACTAAGGCACATCTTTGGCTCTGATGAAGCCTTCAGTAAAGTCAATT 240
QY 181 AATTACCGGACTGAAAATGGGCTGCTCTACTTCTATTTATGTTGATTTGTTGGAGGCAAG 240
Db 241 AATTACCGGACAGAGCGTGGCCTGCTCCCTGCTACACCTCTGCTGTCTGTGGCGGCAAC 300
QY 241 AAATCAATATTCGAACTCTTATGTTGAAAGGGCTCCGCCATCTCGACTGACAAAGAAAT 300
Db 301 AAGTCAATATTCGTCCTCTTATGTTAAAGGGCTCCGTCATCCAGACTGACGAGAAAT 360
QY 301 GGATTTACAGCCTTGAATTTAGCAGTTTACAGGATTAATGAGAAATGATCACTTCTCTG 360
Db 361 GGGTTCCAGCTCTGCACTGGCCGCTTTTAAAGGACAGCCGAACTTATCACTTCACTG 420
QY 361 CTTACAGCTGAGCTGATATACAGAGGTTGGATAGCGGTGGCTCACTGCCCCCTCCATATT 420
Db 421 TTGCACAGCGAGCAGATGTTTACAGAGTGGATAGCGGTGGCTCACAGCCCTCCACATA 480
QY 421 GCTCAATAGCTGCCCACTTAGAGGCTGTGATGCTGTTGCAACATGAGGACTAATGTC 480
Db 481 GCTCAATAGCTGCCCACTTAGAGGCTGTGCAAGGCTGCAAGGCTGCTCTACAAATGGGCGCAATGTC 540
QY 481 AATATTCAGATCAGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 540
Db 541 AATGTTCAAGATCAGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 600
QY 541 CAGTAACTCGCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 600
```

Db 601 CAGTAAACAGTGTCTCTTTTGAAGTTTGGTGTGATGCTCAATGTAAAGCGGTGAAGTTGGG 660  
Qy 601 GATAGACCCCTCCACCTAGCATCTGCAAAAGGATTCTTGAATTAATTCGCAAACTCTTGATG 660  
Db 661 GACAGGCTCTGCACTGCGCTCTGCAAGGCTCTTCAACATTTGTGAAGTCTCTGTA 720  
Qy 661 GAGAGGAGCAAGACAGATGTGATGCTCAAGATATGAAGACCATGTCCCACTCAT 720  
Db 721 GAAGAAGGAGCAAGACAGATGTGAAGCTCAGGCAATGAAGACCAAGTCCCTCTGCAC 780  
Qy 721 TTCTGTTCTCGATTGTCGACCAACATGATAGTAAAGTATCTCTGCAAAAGTGAATTGGAA 780  
Db 781 TTCTGTTCTCGATTGTCGACCAACATGATAGTAAAGTATCTCTGCAAAAGTGAATTGGAA 840  
Qy 781 GTTCAACCTCATGTTGTTAAATATCTATGAGATACCCCTTACCTGGCATGTCTACAT 840  
Db 841 GTCCAGCTCAGCTCATTAACATCTATGTTGACACTCTCTTTGACCTGGCATGTCTACAT 900  
Qy 841 GGCATAATTTGAAGTTCGCAAGGAAATCATCCAAATATCAGGAACAGAAAGTCTGACTAAG 900  
Db 901 GGAATTTTGAAGTTCGCAAGGAAATTTGTCAGGTAAACAGGAACTGAAAGTCTGACTAAG 960  
Qy 901 GAAAAATCTTCAGTGAACACAGCTTTTCAATAGTGTGATCTATGGAAGACATTTGAC 960  
Db 961 GAAAAATCTTCAGTGAACACAGCTTTTCAATAGTGTGATCTATGGAAGACATTTGAC 1020  
Qy 961 CTAGTCAAAATTTCTTGTGATGATGATGATCAATAACATCAACCAAGCAAGGAGGATGG 1020  
Db 1021 CTGGTCAAAATTTCTTGTGATGATGATGATCAATAACATCAACCAAGCAAGGAGATGG 1080  
Qy 1021 CACACTGGATTACACTCTGCTTCTACACAGCTTCACATTCGCTGGTTTCAGTTCTTACTG 1080  
Db 1081 CACACAGGATTGCACTCTGCTTCTACACAGGCTATATCCGCTGGTTTCAGTTCTTACTT 1140  
Qy 1081 GATAATGAGCTGATATGATATCTAGTGGCTTGTGATCCAGCAGTCTAGTGGTGAAGAA 1140  
Db 1141 GATAATGAGCTGATATGATATCTAGTGGCTTGTGATCCAGCAGTCTAGTGGTGAAGAA 1200  
Qy 1141 GATGAGCAGACATGTTTGTATGGGCTTATGAAAGGSCATGATGCCATTTGTCACTC 1200  
Db 1201 GATGAGCAGACATGTTTGTATGGGCTTATGAGAAAGGACATGATGCCATTTGTACACTC 1260  
Qy 1201 CTGAAGCATTTAAGAGACCAAGATGAATGTCCTGTAAATATTTCTCAGGCTCGA 1260  
Db 1261 CTGAAGCATTTAAGAGACCAAGATGAATGTCCTGTAAATATTTCTCAGGCTCGA 1320  
Qy 1261 GGAGATGGCTCTATGCTGTGTTCCATCACCTTGGGAGATTAAGAGCATGACAAA 1320  
Db 1321 GGAGATGGCTCTATGCTGTGTTCCCTTCCCTTGGGCAAGATTAAGAGCATGACAAA 1380  
Qy 1321 GAGAAGCAGATATTTCTCTCTTAAGAGCTGGATTCCTTCAATTTCCATCTTCAGCTC 1380  
Db 1381 GAGAAGCAGATATTTCTCTCTTAAGAGCTGGATTCCTTCAATTTCCATCTTCAGCTC 1440  
Qy 1381 TCAGAAATGAGTTCCATGAGATATTTGGCTCAGTTCTTTGGGAAAGTATTAAGGA 1440  
Db 1441 TCCGAAATCCAGTTCCAGAGATTAATCGGCTGGGTTCTTTTGGGAAAGTATTAAGGG 1500  
Qy 1441 CGATGCAAAATAAATAGTGGCTATAAAGCTTATCGAGCAATACCTTACTGCTCCCAAG 1500  
Db 1501 CGATGCAAAATAAATAGTGGCTATAAAGCTTATCGAGCAATACCTTACTGCTCCCAAG 1560  
Qy 1501 TCAGATGTGATATGTTTCCGAGAGGTGTCATTTCTTCCAGCTGTCATCATCTCCCTGC 1560  
Db 1561 TCAGACGTGATATGTTTCCGAGAGGTGTCATTTCTTCCAGCTGTCATCATCTCCCTGC 1620  
Qy 1561 GTAATTCAGTTTGGGTGCTGCTGATGATCCAGCAGTTTCCCAATTTCTCACTCA 1620  
Db 1621 GTGTTTCAGTTTGGGTGCTGCTGATGATCCAGCAGTTTCCCAATTTCTCACTCA 1680  
Qy 1621 TACATATCAGGGGTTCTCTGTTCTCCCTCTTCATGAGCAGAGAGGATTTCTGATTG 1680  
Db 1681 TACATTTTCAGAGGCTCCCTGTTCTCCCTGTTCTCATGAAACAGAGAGATTTCTGACTTG 1740

Qy 1681 CAGTCTAAATTAATTAATTTGCAGTAGATGTTGCCAAAGCATGGAGTACCTTCAAACTG 1740  
Db 1741 CAGTCTAAATTAATTAATTTGCAGTAGATGTTGCCAAAGCATGGAGTACCTTCAAACTG 1800  
Qy 1741 ACACAGCCAAATTAATATCATCGTGAACAGTTCACAAATTTCTCTATGAGGATGG 1800  
Db 1801 ACCAGCCAAATTAATATCATCGTGAACAGTTCACAAATTTCTCTATGAGGATGG 1860  
Qy 1801 CATGCTGTGGTGGAGATTTTGGAGATTCAGATTTTACAGATCTCTGGATGAAGCAAC 1860  
Db 1861 CATGCTGTGGTGGAGATTTTGGAGATTCAGATTTTCTGAGTCCCTGGATGAAGCAAC 1920  
Qy 1861 ATGACAAAAACACCTCGGAAACCTCGTTGGATGGCTCTGAGTGTTCACGAGTGCACT 1920  
Db 1921 ATGACAAAAACACCTCGGAAACCTCGTTGGATGGCTCTGAGTGTTCACGAGTGCACT 1980  
Qy 1921 CGGTACACCATCAAAAGCAGATCTTTACAGTATGCTGTGTGTGGGAAATTTCTCACT 1980  
Db 1981 AGATACACCATCAAAAGCAGATCTTTACAGTATGCTGTGTGTGGGAGCTCTCACT 2040  
Qy 1981 GCGGAAATTCATTCGCTCATCTCAAGCCAGCGGCTGGGAGCAGACATGGCTTACCAC 2040  
Db 2041 GCGGAAATTCATTCGCTCATCTCAAGCCAGCGGCTGGGAGCAGACATGGCTTACCAC 2100  
Qy 2041 CACATCAGACCTCCCATTTGGCTATTTCCATTTCCAAAGCCCATATCATCTCTCTGATAGA 2100  
Db 2101 CACATCAGACCTCCCATTTGGCTATTTCCATTTCCAAAGCCCATATCATCTCTCTGATAGA 2160  
Qy 2101 GCGTGAACGATGCTCTGAAGAGACCCGAAATTTCTGAAGTGTCTGAAGTTAGAA 2160  
Db 2161 GCGTGAACGATGCTCTGAAGAGACCCGAAATTTCTGAAGTGTCTGAAGTTAGAA 2220  
Qy 2161 GAGTGTCTCTGAAACATTTGAGCTGATGCTCTGATCAAGTAAACAGCAGTGGGTCTCTC 2220  
Db 2221 GAGTGTCTCTGAAACATTTGAGCTGATGCTCTGATCAAGTAAACAGCAGTGGGTCTCTC 2280  
Qy 2221 TCACCTTCTCTTCTTCTGATTTGCTGTAACCCGGGAGACCTGGCCGAGTCAATG 2280  
Db 2281 TCACCTTCTCTTCTTCTGATTTGCTGTAACCCGGGAGACCTGGCCGAGTCAATG 2340  
Qy 2281 GCAGCATTAAGAGTGGTTTCCGAAATGGAATGCTCTAAATGCAAGTCTCTATCTCTCT 2340  
Db 2341 GCAGCATTAAGAGTGGTTTCCGAAATGGAATGCTCTAAATGCAAGTCTCTATCTCTCT 2400  
Qy 2341 TTGTCCAAAGTCTGGCAATATTTCTCTCAAGTCTGTCTTTGGAGAGATGAAGA 2400  
Db 2401 TGTTCCAAAGTCTGGCAATATTTCTCTCAAGTCTGTCTTTGGAGAGATGAAGA 2460  
Qy 2401 AGTCTTCAATACACCCCATTTGACAAATATGCTATGATCCGATCCCATGAGCTCAATG 2460  
Db 2461 AGCACCAGTATTTCAACTGTTCACAAATACGGCTATGTCTGTATCCCATGAGCTGAG 2520  
Qy 2461 CATTTTCATTTCTTGGCAATATGATAGCAGCTTTTGAGCAGCAGC 2505  
Db 2521 CACCTTCACTCCCGCCCAAGCAGCAGCACTTTTGAGCAGCAAC 2565

## RESULT 10

US-09-947-199A-7

; Sequence 7, Application US/09947199A

; Patent No. 6660490

; GENERAL INFORMATION:

; APPLICANT: Raju, Jeyaseelan

; TITLE OF INVENTION: NOVEL CARP PROTEIN AND NUCLEIC ACID MOLECULES AND USES

; FILE OF INVENTION: THEREFOR

; FILE REFERENCE: MNI-068CP2

; CURRENT APPLICATION NUMBER: US/09/947,199A

; CURRENT FILING DATE: 2001-09-05

; PRIOR APPLICATION NUMBER: US 60/111,938

; PRIOR FILING DATE: 1998-12-11

; PRIOR APPLICATION NUMBER: US 09/291,839

; PRIOR FILING DATE: 1999-04-14

;; PRIOR APPLICATION NUMBER: US 09/458,457  
;; PRIOR FILING DATE: 1999-12-10  
;; NUMBER OF SEQ ID NOS: 9  
;; SOFTWARE: PatentIn Ver. 2.0  
;; SEQ ID NO: 7  
;; LENGTH: 3026  
;; TYPE: DNA  
;; ORGANISM: Rattus norvegicus  
;; FEATURE:  
;; NAME/KEY: CDS  
;; LOCATION: (61)..(2565)  
US-09-947-199A-7

Query Match 75.5%; Score 1892.2; DB 4; Length 3026;  
Best Local Similarity 84.7%; Pred. No. 0;  
Matches 2122; Conservative 0; Mismatches 383; Indels 0; Gaps 0;

QY	1	ATGGGAAATTTAATCTAGACCAACCCAACTTGTACTGATGATGGAAGAAAGTC	60
DB	61	ATGGGAAATTTAATCTAGACCAACCCAACTTGTACTGATGATGGAAGAAAGTC	120
QY	61	AGTGAATCATATGTTATCAATAGAAAGATTAGAGATGACCTGCAGATCAAGGAAAAA	120
DB	121	AGTGAATCTTACGCTATTATCATAGAAAGGCTGGAGGATAACCTGCAGATCAAGAAAT	180
QY	121	GAAGTGAAGAACTAAGGAAATATTTGGCTCTGATGAAGCCTTCAGTAAGTCAATTTA	180
DB	181	GAATTTCAAGAACTAAGGAACTATTTGGCTCTGATGAAGCCTTCAGTGAAGTCAGTTTA	240
QY	181	AATTACCCGACCTGAAATAGGGCTCTCTACTTCAATTTATGTTGCAATTTGTGAGGCAAG	240
DB	241	AATTACCCGACCTGAAATAGGGCTCTCTACTTCAATTTATGTTGCAATTTGTGAGGCAAG	300
QY	241	AAATCACATATTCGAATCTTTATGTTGAAAGGGCTCGGCCATCTCGACTGACAGAAAT	300
DB	301	AAATCACATATTCGAATCTTTATGTTGAAAGGGCTCGGCCATCTCGACTGACAGAAAT	360
QY	301	GGATTTACAGCTTGCATTTAGCAGTTTACAAAGGATATGCAAGATATGATGATCTCTCTG	360
DB	361	GGATTTACAGCTTGCATTTAGCAGTTTACAAAGGATATGCAAGATATGATGATCTCTCTG	420
QY	361	CTTACAGTGGAGCTGATATACAGCAGGTTGGATACGGTGGCTCTACCTCCCTCCATATT	420
DB	421	TTGCACACGGAGCAGATGTTTACAGAGTGGGATACGGTGGCTCTACAGCCCTCCACATA	480
QY	421	GCTCAATAGCTGGACACCCAGAGGCTGCAGAGTGGCTGTCTACAAATGGGGCCAAATGTG	540
DB	481	AATATTCAAGATGCAAGTTTCTTCACTCCATTTGATATTCAGCGTACTATGGACATGAA	540
QY	541	AATGTTCAAGATGCGCTCTTCAACCCACTGCACATTCGACCTACTATGGGACAGAG	600
DB	601	CAGGTAACCAAGTCTTTTGAAGTTTGGTCTGATGTCATTAAGCGGTGAAGTTGGG	660
QY	601	GATAGACCCCTCCACCTAGCATCTGCAAAAGGATTTCTTGAATATTCGAAATCTTTGATG	660
DB	661	GACAGGCTCTGCACCTGGCTCTGCAAGGGCTTCTTCAATTTGATGAACTCTCTGGTA	720
QY	661	GAAAGGACGCAAGCAGATGTAATGCTCAAGATATGAAGACCATGTCCCACTCCAT	720
DB	721	GAAGAGGGAGCAAGCAGATGTAACGCTCAGGCAATGAAGACCATGTCCCTCTGACAC	780
QY	721	TTCTGTTCTCATTTGACACCACTGATATAGTTAAGTATCTGCTGCAAGATGATTTGGA	780
DB	781	TTCTGTTCTCATTTGACACCACTGATATAGTTAAGTATCTGCTGCAAGATGATTTGGA	840
QY	781	GTTCAACCTCATGTTGTTAATATCTATGGAGATACCCCTTACACCTGGCATGCTACAAAT	840
DB	841	GTCCAGGCTCAGCTCATTAACATCTATGTTGACACTCTCTTTGACCTGTCATGCTTCAAT	900

QY	841	GGCAAAATTTGAAGTTGCCAAGGAAATCATCCAAATATCAGGAACAGAAAGTCTGACTAAG	900
DB	901	GGAAATTTTGAAGTTGCCAAGGAAATTTGCCAGGTAACAGGAACAGAAAGTCTGACTAAG	960
QY	901	GAAAAATCTTTCAGTGAACAGCTTTTTCATAGTGTCTTGTACCTATGCGCAAGAGCATTTGAC	960
DB	961	GAAAAATCTTTCAGTGAACAGCTTTTTCATAGTGTCTTGTACCTATGCGCAAGAGCATTTGAC	1020
QY	961	CTAGTCAAAATTTCTTCTTGTATCAGAATGCTATAAAATCAACCAACCAAGGAGGATGGG	1020
DB	1021	CTAGTCAAAATTTCTTCTTGTATCAGAATGCTATAAAATCAACCAACCAAGGAGGATGGG	1080
QY	1021	CACACTGGATTAACACTCTGCTCTACACCGGTACATTCGCTGGCTGGTTCAGTTCCTACTG	1080
DB	1081	CACACTGGATTAACACTCTGCTCTACACCGGTACATTCGCTGGCTGGTTCAGTTCCTACTG	1140
QY	1081	GATAATGAGCTGATATGATCTAGTGGCTTGTGATCCCGGATGATCCCGGATGATCCCGGAT	1140
DB	1141	GATAATGAGCTGATATGATCTAGTGGCTTGTGATCCCGGATGATCCCGGATGATCCCGGAT	1200
QY	1141	GATGAGCAGACATGTTTGTATGTTGGCTTATGAAAAAGGCGCATGATGCCATTTGCACATC	1200
DB	1201	GATGAGCAGACATGTTTGTATGTTGGCTTATGAAAAAGGCGCATGATGCCATTTGCACATC	1260
QY	1201	CTGAAGCATTATTAAGAGACCAAGATGAATTCCTTGGGAGATTTAAAGAGCATGACAAAA	1260
DB	1261	CTGAAGCATTATTAAGAGACCAAGATGAATTCCTTGGGAGATTTAAAGAGCATGACAAAA	1320
QY	1261	GGAGATGGCTCTTATGTTCTGTTCCATCACCTTGGGAGATTTAAAGAGCATGACAAAA	1320
DB	1321	GGAGATGGCTCTTATGTTCTGTTCCATCACCTTGGGAGATTTAAAGAGCATGACAAAA	1380
QY	1321	GAGAAGCAGATATTTCTCTCCTAGAGCTGATGTCCTTCAATTTCCATCTTCAGCTC	1380
DB	1381	GAGAAGCAGATATTTCTCTCCTAGAGCTGATGTCCTTCAATTTCCATCTTCAGCTC	1440
QY	1381	TCAGAAATTCAGTTCATGAGATTTATGCTCAGGTTCTTTGGGAGATTTAAAGAGCATGACAAAA	1440
DB	1441	TCAGAAATTCAGTTCATGAGATTTATGCTCAGGTTCTTTGGGAGATTTAAAGAGCATGACAAAA	1500
QY	1441	CGATGCAAGAAATTAATAGTGGCTATAAAAGCTTATCGAGCAATACCTTCTGCTCCAG	1500
DB	1501	CGATGCAAGAAATTAATAGTGGCTATAAAAGCTTATCGAGCAATACCTTCTGCTCCAG	1560
QY	1501	TCAGATGCTGATATGTTTTCGAGAGGCTCCATTTCTCTGCGAGCTCAATCATCTCCCTGC	1560
DB	1561	TCAGATGCTGATATGTTTTCGAGAGGCTCCATTTCTCTGCGAGCTCAATCATCTCCCTGC	1620
QY	1561	GTAATTCAGTTTGGGCTCTTGTGTAATGATCCCGAGCTTTGCGCATTTGCTCACTCAA	1620
DB	1621	GTAATTCAGTTTGGGCTCTTGTGTAATGATCCCGAGCTTTGCGCATTTGCTCACTCAA	1680
QY	1621	TACATATCAGGGGTTCTCTGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1680
DB	1681	TACATATCAGGGGTTCTCTGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1740
QY	1681	CAGTCTAAATTAATTTAGTGGCTAGATGTTGCCAAGGATGAGTACCTTCAACACCTG	1740
DB	1741	CAGTCTAAATTAATTTAGTGGCTAGATGTTGCCAAGGATGAGTACCTTCAACACCTG	1800
QY	1741	ACACAGCAATTTATAGCTGATGTAAGTGAACAGTCAATATTTCTCTCTCTCTCTCTCTCTCT	1800
DB	1801	ACACAGCAATTTATAGCTGATGTAAGTGAACAGTCAATATTTCTCTCTCTCTCTCTCTCTCT	1860
QY	1801	CATGCTGTGTGTCAGATTTTGGAGATCAAGATTTCTACAGTCTCTCTCTCTCTCTCTCTCT	1860
DB	1861	CATGCTGTGTGTCAGATTTTGGAGATCAAGATTTCTACAGTCTCTCTCTCTCTCTCTCTCT	1920
QY	1861	ATGACAAAAACACTCTGGAACTCTGTTGGATGGCTCTGAGGTTGTTCCAGGATGCTACT	1920
DB	1921	ATGACAAAAACACTCTGGAACTCTGTTGGATGGCTCTGAGGTTGTTCCAGGATGCTACT	1980
QY	1921	CGGTACACCATCAAGACAGATGCTTTCAGCTATGCTCTGTTGCTGTGTGTGTGTGTGTGTGT	1980

```

Db      1981 AGATACACATCAAGGCTGATGCTTCAAGTACTCCCTGTGTCTGTGGAGCTCCTCACT 2040
Qy      1981 GGGAAATTCATTCGCTCATCTCAAGCAGCGGCTGCGGAGCAGACATGCTTACCAC 2040
Db      2041 GGAAGAAATTCATTCGCTCATCTCAAGCAGCGGCTGCGGAGCAGACATGCTTACCAC 2100
Qy      2041 CACATCAGACCTCCCATGCTGCTATTCATTCCTCCAGCCCATATCATCTCTGCTGATACGA 2100
Db      2101 CACATCAGACCGGCCATCGGCTATTCATTCCTCCAGCCCATATCATCTCTGCTGATACGG 2160
Qy      2101 GGGTGAACCATGTCTGTAAGGAAGACCGCAATTTCTGAAGTTGTTCATGAAGTTGAA 2160
Db      2161 GGGTGAATGCATGTCTGAAGACGACGACAGAGTCTCTGAAGTGGTTAGCAAACTGGAG 2220
Qy      2161 GAGTGTCTCTGCAACATGAGCTGATGCTCTCTGCAATCAAGTAAACAGCAGTGGTCTCTC 2220
Db      2221 GAGTGGCTATGCAATGTGGAGCTCATGTCTTCAGCATCAAGTAACAGCAGTGGTCTCTG 2280
Qy      2221 TCACCTTCTTCTTCTCTGATTCCTGATTCCTGGTGAACCGGGGAGGACCTGGCGGAGTCATGTG 2280
Db      2281 TCACCTTCTTCTTCTCCGATTCCTGTGAGCGGGGAGGCTGGCGGAGCCACGTG 2340
Qy      2281 GCAGCATTAAGAGTGTCTTTCGAATGGAATATGCTCTAAATGCAAGTCTCTATGCTGCT 2340
Db      2341 GCAGCCTTACGGAGCGGTTTTTGAGTTGGAGTATGCCCTAAATGCAAGTCTCTATGCTGG 2400
Qy      2341 TTGTCCCAAGTGTCTGACAAATATCTCTCAAGGTCTGTCTTGGAGGAGATGAAAGA 2400
Db      2401 TGGTCCCAAGTGTGGAACACACTCTAATCCGGGCTGTCTTTGGAGGAGATGAATAGG 2460
Qy      2401 AGTCTTCAATACACCCATGTAACAATATGGCTATGATTCGATCCCATGAGCTCAATG 2460
Db      2461 AGCACCCAGTATTCAACTGTGTGAACAATACGGCTATGTCTGATCCCATGAGCCTGACG 2520
Qy      2461 CATTTTCATTTCTGCGCAATAGTACAGCTTTGAGCAGCAGC 2505
Db      2521 CACCTTCACCTCCCGCAAGACGACGCAACTTTGAGCAGCAAC 2565

RESULT 11
US-09-833-381-1149
; Sequence 1149, Application US/09833381
; Patent No. 6672186
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; PRIOR FILING DATE: 2001-04-11
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: Fast-Seq for Windows Version 3.0
; SEQ ID NO 1149
; LENGTH: 616
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-833-381-1149

Query Match      17.1%; Score 427.4; DB 4; Length 616;
Best Local Similarity 81.8%; Pred. No. 3.1e-124;
Matches 505; Conservative 0; Mismatches 111; Indels 1; Gaps 1;

Qy      291 GACAAGAAATGGATTTACAGCTTGCACTTTAGCAGTTTACAGGATAATGCAAGATTTGAT 350
Db      1 GACGAGAAATGGGTTTCCAGCTCTGCACTTGGCGGCTTTTACAGGACAGCCCGCACTTAT 60
Qy      351 CACTTCTCTGCTTCCAGCTGGAGCTGATATACAGAGGTTGGATACGGTGGCTCACTGC 410
Db      61 CACTTCACTGTTGCACAGCGGACGAGATGTTCAGCAAGTGGGATACGGTGGCTCACAG 120
Qy      411 CCTCCATATTCTACAATAGCTGGCCACCTAGAGGCTGTGATGTGCTGTTGCAACATGG 470
```

```

Db      121 CCTCC-CATAGCTGCAATAGCTGGACACCCAGAGGCTGCAGAAGTGTCTGTCAACATGG 179
Qy      471 AGCTAATGTCATATTCAGATGCGAGTTTTTTTCACTCCATGCAATATGCGAGCTACTA 530
Db      180 GGGCAACGTTGAATGTTCAAGATGCGGTCTTCTTACCCCACTGCACATGCGAGCCTACTA 239
Qy      531 TGGACATGAACAGGTAACCTCGCTCTTCTTTTGAATTTGGTGTGATGTAAATGTAAGTGG 590
Db      240 TGGGACAGCAGCAGGTAACCCAGTGTCTTTTGAAGTTTGGTGTGATGTCAATGTAAGCGG 299
Qy      591 TGAAGTTGAGATAGACCCCTCCACTAGCATCTCTGCAAAAGGATTTCTTGAATATTTGCAAA 650
Db      300 TGAAGTTGGGGACAGGCTCTGCACCTGGCCTCTGCAAGGGCTTCTTCAACATTTGTGAA 359
Qy      651 ACTCTTGATGGGAAGGAGCAGCAAGCAGATGTAATGCTCAAGATAATGAAGACCATGT 710
Db      360 ACTCTTGATGAAGAGGAGCAAGCAGATGTGAACGCTCAGGCAATGAGACCATGT 419
Qy      711 CCCACTCCATTTCTGTTCTCGATTTGGACACCATGATATAGTTAAGTATCTGCTGCAAAAG 770
Db      420 CCTCTGCACTTCTGTTCTCGATTTGGACACCAATATAGTGAGCTACCTGCTCCAGAG 479
Qy      771 TGATTTGGAAGTTCACCTCATGTTGTTAATATCTATGGAGATACCCCTTACACCTGGC 830
Db      480 TGACTTAGAGGTCAGGCTCAGCTCATTAACATCTATGTGACACTCTCTTTGACCTGGC 539
Qy      831 ATGCTACAATGGCAAAATTTGAAGTTGCCAAGGAAATCATCCAAATATCAGGAACAGAAAG 890
Db      540 ATGCTACAATGGAAATTTTGAAGTTGCCAAGGAAATTTGTCAGGTAACAGGAACCTGAAAG 599
Qy      891 TCTGACTAAGGAAAAACA 907
Db      600 TCTGACTAAGGAAAAACA 616

RESULT 12
US-09-833-381-1150
; Sequence 1150, Application US/09833381
; Patent No. 6672186
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; PRIOR FILING DATE: 2001-04-11
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: Fast-Seq for Windows Version 3.0
; SEQ ID NO 1150
; LENGTH: 736
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; LOCATION: (1)-(736)
; OTHER INFORMATION: n = A,T,C or G
US-09-833-381-1150

Query Match      16.6%; Score 414.6; DB 4; Length 736;
Best Local Similarity 81.2%; Pred. No. 3.8e-120;
Matches 480; Conservative 0; Mismatches 111; Indels 0; Gaps 0;

Qy      265 TTGAAGGGTCCGCCATCTCGACTGACAAGAAATGGATTTACAGCCTTGCAATTAGCA 324
Db      1 TTAAGAGGGTCCGTCATCCAGACTGACGAGAAATGGGTTTCCAGCTCTGCACCTGGCC 60
Qy      325 GTTTACAAGGATAATGCAGAAITGACACTTCTCTGCTTCACAGTGGAGCTGATATACAG 384
Db      61 GTTTACAAGGACAGCCCGGAACTTACCTTCACTTGTTCACAGCGGAGCAGATGTTCAAG 120
Qy      385 CAGTTGGATACGGTGGCCTCACTGCCCTCCATTTGCTACAAATAGCTGGCCACCTAGAG 444
```

Db 121 CAAAGTGGGATACGGTGGCTCCACAGCCCTCCACATAGCTCAATAGCTGGACACCCAGAG 180  
QY 445 GTCGCTGATGCTGCTTGCACATGAGAGCTAATGCAATATTAAGATGAGTGTGTTTTC 504  
Db 181 GCTGCAGAAAGTCTGCTCAACATGGGCAATGTGAATGTTCAAGATGCGCTGCTTTC 240  
QY 505 ACTCCATTGCATATTGACAGCTACTATGACATCAAGAGGTAACTCCCTTCTTTTGAAA 564  
Db 241 ACCCACTGCACATTTGAGCTACTATGGGACAGCAGGTAAACAGTGTCTTTTGAAG 300  
QY 565 TTGTGCTGATGTAATGTAAGTGGTGAAGTTGGAGATAGACCCCTCCACCTAGCATCT 624  
Db 301 TTGTGCTGATGATCAATGTAAGCGTGAAGTTGGGACAGCCCTCTGCACCTGGCCTCT 360  
QY 625 GCAAAAGGATCTTGAATATTGCAAACTCTTGATGGAGAGGACCAAGCAGATGTG 684  
Db 361 GCAAAAGGCTTCTTCAACATTTGTGAACCTCTGTGTAAGAGGAGCAAGCAGATGTG 420  
QY 685 AATGCTCAAGATAATGAAGACCATGTCCCACTCCATTTCTGTTCTCGATTTGACACCAT 744  
Db 421 AACGCTCAGGCAATGAAGACCATGCTCCCTCTGCATTTCTGCTCGATTTGACACCAT 480  
QY 745 GATATAGTATGATCTCTGCTGCAAGTGAATTTGGAAGTTCACCTCATGTTGTTAATATC 804  
Db 481 AATATAGTATGATCTCTGCTGCAAGTGAATTTGGAAGTTCACCTCATGTTGTTAATATC 804  
QY 805 TATGAGATACCCCTTACACTGCGCATGCTACAAATGCAAAATTTGAAGTT 855  
Db 541 TATGATGACATCTCTTGCACCTGGCATGCTACAAATTTGAATTT 591

## RESULT 13

US-09-833-381-1151  
; Sequence 1151, Application US/09833381  
; Patent No. 6672186  
; GENERAL INFORMATION:  
; APPLICANT: Robison, Keith E.  
; TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs  
; FILE REFERENCE: 5800-119  
; CURRENT APPLICATION NUMBER: US/09/833,381  
; PRIOR FILING DATE: 2001-04-11  
; PRIOR APPLICATION NUMBER: 09/516,448  
; PRIOR FILING DATE: 2000-02-29  
; NUMBER OF SEQ ID NOS: 2050  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1151  
; LENGTH: 740  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(740)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-833-381-1151

Query Match 16.5%; Score 413.2; DB 4; Length 740;  
Best Local Similarity 79.0%; Pred. No. 1.1e-119;  
Matches 498; Conservative 0; Mismatches 131; Indels 1; Gaps 1;

QY 1 ATGGGAATTTAAATCTAGACCAACCCAACTTGACTGATGAATGGAGAAAAGTTC 60  
Db 73 ATGGGAATTTAAATCTAGACCAACCAACAGACTTGTCTGATGAATGGAGAAAAGTTC 132  
QY 61 AGTGAATCATATGTTAATCAATATGAAGATTAGACATGACCTGAGATCAAGGAAA 120  
Db 133 AGTGAATCTTACGCTATTATCATGAAGGCTGGAGATACTTGCAGATCAAGAAAAT 192  
QY 121 GAACCTGACAGAACTAAGGAATATATTGGCTCTGATGAAGCCCTTCAGTAAGTCAATTTA 180  
Db 193 GAATTTCAAGAACTAAGGCACATCTTTGGCTCTGATGAAGCCCTTCAGTAAGTCAATTTA 252  
QY 191 AATTACCGCACTGAAAATGGGCTCTCTACTTCATTTATGTTGCAATTTTGTGGAGCAAG 240

Db 253 AATTACCGCAGAGCTGGCTGTCCCTGCTACACCTCTGCTGTCTGTGGCGCAAC 312  
QY 241 AAATCAGATATTCGAACTCTTATGTTGAAAGGCTCCGCCATCTCGACTGACAGAAAT 300  
Db 313 AAGTCAGATATCCGTCGCTTATGTTAAAGGCTCCGTCATCCAGACTGACGAGAAAT 372  
QY 301 GGATTTACAGCTTTGCAATTTAGCAGTTTACAAGGATAATGCAAAATTTGATCACTTCTCTG 360  
Db 373 GGGTTTCAGCTCTGCACTGGCGCTTTTACAAGGACAGCCCGGAATTTATCACTTCACTG 432  
QY 361 CTTTAC-AGTGGAGCTGATATACAGCAGGTTGGATACGGTGGCCTCACTGCCCTCCATAT 419  
Db 433 TTGCACAAAGGAGCAGATGTTTCAAGCAAGTGGGATACGGTGGCCTCACAGCCCTCCACAT 492  
QY 420 TGCTACATAGCTGGCCACCTAGAGGCTGCTGATGCTGTTGCAACATGAGGCTAATGT 479  
Db 493 AGCTGCAATAGCTGGACACCCAGAGGCTGCANAAAGTCTGNTGCAACATGGGGCCCAAGT 552  
QY 480 CAATATTTCAAGATGCACTTTTTTTTCACTCCATTTGCATATTTGAGGCTACTATGGACATGA 539  
Db 553 GAATGTTCAAGATGCGCTCTTCTTCACTCCACTGCACATTTGAGGCTACTATGGGACGA 612  
QY 540 ACAGGTAACCTGCGCTTCTTTTGAATTTGGTCTGATGTAATGTAAGTGGTGAAGTTGG 599  
Db 613 GCAGGTAACCAAGTGTCTTTTGAAGTTTGGGCTGATGTCNATGTAAGCGGTAAAGTTGG 672  
QY 600 AGATAGACCCCTCCACCTAGCATCTGAAA 629  
Db 673 GNACAGGCTTCGCGCCCTGCGCCNCCGNAAA 702

## RESULT 14

US-09-833-381-1147/c  
; Sequence 1147, Application US/09833381  
; Patent No. 6672186  
; GENERAL INFORMATION:  
; APPLICANT: Robison, Keith E.  
; TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs  
; FILE REFERENCE: 5800-119  
; CURRENT APPLICATION NUMBER: US/09/833,381  
; PRIOR FILING DATE: 2001-04-11  
; PRIOR APPLICATION NUMBER: 09/516,448  
; PRIOR FILING DATE: 2000-02-29  
; NUMBER OF SEQ ID NOS: 2050  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1147  
; LENGTH: 304  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(304)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-833-381-1147

Query Match 7.4%; Score 184.2; DB 4; Length 304;  
Best Local Similarity 79.6%; Pred. No. 8.6e-48;  
Matches 242; Conservative 0; Mismatches 59; Indels 3; Gaps 2;

QY 266 TGAAGGGCTCCGCCCATCTCGACTGACAAAGAAATGGATTTACGCC--TTGCAATTTAGC 323  
Db 304 TAAAGGGCTCCGTCCTCCATCCAGACTGACGAGAAATGGGTTTCCAAAGCTTTTGCACTGCC 245  
QY 324 AGTTTACAGGATAATGCAAGATTGATCAC-TTCTCTGCTTCACAGTGGAGCTGATATAC 382  
Db 244 GGTTTACAAGGACAGCCCGGAATTTATCACTTTCTACTGTGTCACAGCGAGCAGATGTC 185  
QY 383 AGCAGGTTGGATACGGTGGCCTCACTGCCCTCCATATTTGCTACAAATAGTGCCACCTAG 442  
Db 184 AGCAAGTGGGATACGGTGGCCTCACAGCCCTCCACATAGCTGCAATAGCTGACACCCAG 125  
QY 443 AGCTGCTGATGCTGTTGCAACATGAGGCTAATGTCATATTCAGATGCGAGCTTTTTT 502

Db 124 AGGCTGAGAAAGTCTGATGCAACATGGGGCCCAAGCTGAATGTTCAAGATCCGCTTCT 65  
Qy 503 TCACTCCATTGCATATTGACGGTACTATGGACATGAACAGGTAACCTGCTTTTGA 562  
Db 64 TCACCCCACTGCACATGCGAGTACTATGGGACAGAGGTAACCAAGTGTCTTTTGA 5  
Qy 563 AATT 566  
Db 4 AGTT 1

RESULT 15

US-09-833-381-1148/c  
; Sequence 1148, Application US/09833381  
; Patent No. 6672186  
; GENERAL INFORMATION:  
; APPLICANT: Robison, Keith E.  
; TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs  
; FILE REFERENCE: 5800-119  
; CURRENT APPLICATION NUMBER: US/09/833,381  
; CURRENT FILING DATE: 2001-04-11  
; PRIOR APPLICATION NUMBER: 09/516,448  
; PRIOR FILING DATE: 2000-02-29  
; NUMBER OF SEQ ID NOS: 2050  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1148  
; LENGTH: 186  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)...(186)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-833-381-1148

Query Match 3.4%; Score 84.4; DB 4; Length 186;  
Best Local Similarity 69.1%; Pred. No. 1.7e-16;  
Matches 112; Conservative 0; Mismatches 50; Indels 0; Gaps 0;  
Qy 405 CACTGCCCTCCATATTGCTACAAATAGCTGGCCACCTAGAGGCTGCTGATGCTGTTCCA 464  
Db 162 CCCTTCCAANAATAGTGCCCAATAAGTTGNACACCCCAANAGGTGCAAAAAGTGTGTTNCA 103  
Qy 465 ACATGGAGCTAATGTCATATTCAAGATGAGTCTTTTCTCACTCCATTCATATTGCAGC 524  
Db 102 ACATGGGGCCAATGTGAATGTTCAAGATGCCGCTCTTCTTCAACCCCACTGCACATTGCAGC 43  
Qy 525 GTACTATGGACATGAACAGGTAACCTGCTTTTGAATT 566  
Db 42 CTACTATGGGCACGAGCAGGTAACCAAGTGTCTTTTGAAGTT 1

Search completed: September 6, 2004, 07:42:53  
Job time : 124.494 secs

This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 5, 2004, 23:38:08 ; Search time 4170.17 Seconds  
(without alignments)  
17938.068 Million cell updates/sec

Title: US-10-626-173-3

Perfect score: 2505

Sequence: 1 atgggaataataatctag.....gcagctttgagacagcgc 2505

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: em\_estba:\*\*

2: em\_esthum:\*\*

3: em\_estlin:\*\*

4: em\_estmu:\*\*

5: em\_estov:\*\*

6: em\_estpl:\*\*

7: em\_estro:\*\*

8: em\_htc:\*\*

9: gb\_est1:\*\*

10: gb\_est2:\*\*

11: gb\_htc:\*\*

12: gb\_est3:\*\*

13: gb\_est4:\*\*

14: gb\_est5:\*\*

15: em\_estfun:\*\*

16: em\_estom:\*\*

17: em\_gss\_hum:\*\*

18: em\_gss\_inv:\*\*

19: em\_gss\_pln:\*\*

20: em\_gss\_vrt:\*\*

21: em\_gss\_fun:\*\*

22: em\_gss\_mam:\*\*

23: em\_gss\_mus:\*\*

24: em\_gss\_pro:\*\*

25: em\_gss\_rtd:\*\*

26: em\_gss\_phg:\*\*

27: em\_gss\_vrl:\*\*

28: gb\_gss1:\*\*

29: gb\_gss2:\*\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	734.6	29.3	876	14	CF551822 AGENCOURT
2	721	28.8	756	12	BG227974 RST41894
3	682.8	27.3	813	12	BG209217 RST28732
4	677.4	27.0	793	12	BG216613 RST16305

5	665	26.5	665	13	BX501623
6	618.6	24.7	902	10	BB897696
7	587.4	23.4	656	13	BX502807
8	557	22.2	638	9	AL599654
9	491	19.6	491	13	BX112488
c 10	410.4	16.4	712	13	BQ622977
11	392.8	15.7	507	10	BF556094
12	382.4	15.3	862	9	AU130402
13	374	14.9	950	14	CF412400
14	358.6	14.3	634	10	BB658710
15	356	14.2	465	13	BX099674
16	355	14.2	1113	14	CD513739
c 17	347	13.9	884	9	AI375137
c 18	345	13.8	648	13	BQ774566
19	333.6	13.3	796	9	AU051569
c 20	330.6	13.2	514	10	BF197082
c 21	329.6	13.2	514	10	BF197116
22	324.4	13.0	667	14	CA589597
23	321.4	12.8	328	13	C03950
24	321.4	12.8	757	12	BG219957
25	321	12.8	820	13	B013846
c 26	317.4	12.7	530	9	AI333762
27	310	12.4	446	12	BG406579
28	306.6	12.2	575	10	AW140835
29	301	12.0	623	10	B8659579
30	301	12.0	643	10	B8662065
31	301	12.0	705	11	AK052882
32	301	12.0	741	11	AK084817
33	286.8	11.1	673	10	BB480482
34	278	11.1	662	10	BB516995
35	269.2	10.7	574	13	B0447828
36	259.8	10.4	693	12	BF547671
37	255.2	10.2	699	13	B0358849
c 38	254.2	10.1	519	10	BE501083
39	250.8	10.0	535	14	CB718149
40	247.6	9.9	834	13	B0386757
41	238	9.5	531	14	CB719013
c 42	237.4	9.5	452	9	AI991354
c 43	232.2	9.3	413	9	AI040439
c 44	225.2	9.0	384	13	BY086637
c 45	206.6	8.2	512	10	AW631150

#### ALIGNMENTS

RESULT 1  
CF551822  
LOCUS  
DEFINITION AGENCOURT 15595443 NIH\_MGC 183 Homo sapiens linear EST 22-SEP-2003  
IMAGE:30530169 5', mRNA sequence.  
ACCESSION CF551822  
VERSION CF551822.1 GI:34888656  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 876)  
NIH-MGC <http://mgc.ncbi.nlm.nih.gov/>  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Daniela S. Gerhard, Ph.D.  
Office of Cancer Genomics  
National Cancer Institute / NIH  
Bldg. 31 Rm10A07 Bethesda, MD 20892  
Email: [cgabbs@mail.nih.gov](mailto:cgabbs@mail.nih.gov)  
Tissue Procurement: Dr. Michael Brownstein  
CDNA Library Preparation: Invitrogen Corp  
DNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)  
DNA sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at:

http://image.llnl.gov

Plate: NDAM618 row: 0 column: 10

High quality sequence start: 15

High quality sequence stop: 670.

## FEATURES

source

Location/Qualifiers

1..876

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:30530169"

/lab\_host="DH10B-TonA (T1 and T5 phage resistant)"

/clone\_lib="NH MGC 183"

/note="Organ: Pooled muscle (cardiac and skeletal);"

Vector: pCMV-SpORf6.1; Site\_1: EcoRV (destroyed); Site\_2:

Not; Library is oligo-dr primed and directionally cloned

(EcoRV site is destroyed upon cloning). Average insert

size 1.7. Library was constructed by Invitrogen."

## ORIGIN

Query Match 29.3%; Score 734.6; DB 14; Length 876;  
Best Local Similarity 97.6%; Pred. No. 2.4e-189;  
Matches 830; Conservative 0; Mismatches 10; Indels 10; Gaps 8;

QY 352 ACTTCTCTGCTTACAGTGGAGCTGATATACAGCAGGTGGATACGGTGGCCCTCACTGCC 411  
DB 28 ACTTCTCTGCTTACAGTGGAGCTGATATACAGCAGGTGGATACGGTGGCCCTCACTGCC 86  
QY 412 CTCATATGTCTACATAGCTGGCCACTAGAGCTGCTGATGTGCTGTGGCAATGGA 471  
DB 87 CTCATATGTCTACATAGCTGGCCACTAGAGCTGCTGATGTGCTGTGGCAATGGA 146  
QY 472 GCTATGTCTACATAGCTGGCCACTAGAGCTGCTGATGTGCTGTGGCAATGGA 531  
DB 147 GCTATGTCTACATAGCTGGCCACTAGAGCTGCTGATGTGCTGTGGCAATGGA 206  
QY 532 GGACATGAACAGGTAACTCGCTTCTTTTGAATTTGGTCTGATGTAATGTAAGTGT 591  
DB 207 GGACATGAACAGGTAACTCGCTTCTTTTGAATTTGGTCTGATGTAATGTAAGTGT 266  
QY 592 GAAGTGGAGATAGACCCCTCCACTAGCATCTGCAAAAGGATCTTTGAATTTGCAAAA 651  
DB 267 GAAGTGGAGATAGACCCCTCCACTAGCATCTGCAAAAGGATCTTTGAATTTGCAAAA 326  
QY 652 CTCCTGATGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 711  
DB 327 CTCCTGATGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 386  
QY 712 CCACATCAATTTCTGTTCTGATTTGGACACCATGATATAGTAAATGTAAGTCTGCTGCAAGT 771  
DB 387 CCACATCAATTTCTGTTCTGATTTGGACACCATGATATAGTAAATGTAAGTCTGCTGCAAGT 446  
QY 772 GATTTGGAAGTTCACCTCATGTTGTTAAATCTATGGAGATACCCCTTTACACCTGGCA 831  
DB 447 GATTTGGAAGTTCACCTCATGTTGTTAAATCTATGGAGATACCCCTTTACACCTGGCA 506  
QY 832 TGCTCAATGGAATTTGAGTTGCGCAAGAAATCATCCAAATATCAGAAACAGAAAGT 891  
DB 507 TGCTCAATGGAATTTGAGTTGCGCAAGAAATCATCCAAATATCAGAAACAGAAAGT 566  
QY 892 CTGACTAAGGAAACATCTTCAGTGAACACAGCTTTTCATAGTCTGTTTACCTATGCAAG 951  
DB 567 CTGACTAAGGAAACATCTTCAGTGAACACAGCTTTTCATAGTCTGTTTACCTATGCAAG 626  
QY 952 AGCATTGACCTAGTCAAAATTTCTTTGATCAGAAATGTCATAACATCAACACCAAGG- 1010  
DB 627 AGCATTGACCTAGTCAAAATTTCTTTGATCAGAAATGTCATAACATCAACACCAAGG 686  
QY 1011 AAGGATGGGACACT-GGATTACACTCT-GCTTGTACACCGGTCACTTCG-CCTGCT 1067  
DB 687 AAGGATGGGACACTNGGATTACACTCTGCTTGTACACCGGTCACTTCG-CCTGCT 746  
QY 1068 TCAGTCTTCTACTGGATAATGGAGCTGATGATGATCTAGT-GGCTTGTGATCCAGGAG- 1125

Db 747 TCAGTCTTCTACTGGATAATGGAGCTGATATGATCTAGTGGCTTGTGATCCACGACG 806  
QY 1126 TCTAGTGGTGAAGAGATGAGCAGCATGTTT---CATGTGGCTTATGAAAAAGGGCAT 1182  
Db 807 TCTAGTGGTGAAGAGATGAGCAGCATGTTTGTGATGTTGGGCTTATGAAAAAGGGCAT 866  
QY 1183 GATGCCATTG 1192  
Db 867 GATGCCATTG 876

RESULT 2  
BG227974 756 bp mRNA linear EST 21-APR-2001  
LOCUS BG227974 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.  
DEFINITION BG227974  
ACCESSION BG227974  
VERSION BG227974.1 GI:13748097  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE 1 (bases 1 to 756)  
AUTHORS Harrington, J.J., Sherf, B., Rundlett, S., Jackson, P.D., Perry, R.,  
Cain, S., Leventhal, C., Thornton, M., Ramachandran, R.,  
Wittington, J., Lerner, L., Costanzo, D., McElligott, K., Booser, S.,  
Mays, R., Smith, E., Veloso, N., Klika, A., Hess, J., Cothren, K., Lo, K.,  
Offenbacher, J., Danzig, J., and Ducar, M.  
TITLE Creation of genome-wide protein expression libraries using random  
activation of gene expression  
JOURNAL Nat. Biotechnol. 19 (5), 440-445 (2001)  
MEDLINE 21227151  
PUBMED 11329013  
COMMENT Contact: Scott J. Cain  
Athersys, Inc.  
3201 Carnegie Ave, Cleveland, OH 44115, USA  
Tel: 216 431 9900  
Fax: 216 361 9596  
Email: scain@atersys.com  
High quality sequence stop: 520.  
FEATURES  
source  
1..756  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/cell\_line="HT1080"  
/clone\_lib="Athersys RAGE Library"  
/note="See 'Creation of Genome-wide Protein Expression  
Libraries using Random Activation of Gene Expression',  
Nature Biotechnology, in press. Note that even though the  
cell type indicated is HT1080, since a random activation  
method was used, these sequence tags are not necessarily  
expressed in HT1080 under normal circumstances."

## ORIGIN

Query Match 28.8%; Score 721; DB 12; Length 756;  
Best Local Similarity 97.3%; Pred. No. 1.2e-185;  
Matches 733; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 546 AACTCGCTCTCTTTTGAATTTGGTCTGCTGATGTAATGTAAGTGTGGAAGTTGGAGATAG 605  
DB 1 AACTCGCTCTCTTTTGAATTTGGTCTGCTGATGTAATGTAAGTGTGGAAGTTGGAGATAG 60  
QY 606 ACCCTCCACTAGCATCTGCAAAAGGATCTTGAATTTGCAAACTTTGATGGAAGA 665  
DB 61 ACCCTCCACTAGCATCTGCAAAAGGATCTTGAATTTGCAAACTTTGATGGAAGA 120  
QY 666 AGGAGCAAGACAGATGTGAATGCTCAAGATAATGAAGACCATGTCCCACTCCATTTCTG 725  
DB 121 AGGAGCAAGACAGATGTGAATGCTCAAGATAATGAAGACCATGTCCCACTCCATTTCTG 180  
QY 726 TTCTCGATTTGGACACCATGATATGTTAAGTATCTGCTGCAAGTATTTGGAAGTTCA 785

```
Db 181 TTCTTTATTGGACACCATGATATAGTTAAGTATCTGCTGCAAGTGAATTTGGAAGTTCA 240
Qy 786 ACCTCATGTTGTTAATATCTATGGAGATACCCCTTTACACCTGGCATCTACAATGGCAA 845
Db 241 ACCTCATGTTGTTAATATCTATGGAGATACCCCTTTACACCTGGCATCTACAATGGCAA 300
Qy 846 ATTGGAAGTTGCCAAGGAATCATCAATATTCAGAAATCAGAAAGTCTGACTAAGGAAA 905
Db 301 ATTGGAAGTTGCCAAGGAATCATCAATATTCAGAAATCAGAAAGTCTGACTAAGGAAA 360
Qy 906 CATCTTTCACTGAAACAGCTTTTCATAGTCTGTTGACCTATGCAAGAGCATTTGACCTAGT 965
Db 361 CATCTTCAGTGAACAGCTTTTCATAGTCTGTTGACCTATGCAAGAGCATTTGACCTAGT 420
Qy 966 CAAATTTCTTTGATCAGAAATCATCAATATTCAGAAATCAGAAAGTCTGACTAAGGAAA 1025
Db 421 CAAATTTCTTTGATCAGAAATCATCAATATTCAGAAATCAGAAAGTCTGACTAAGGAAA 480
Qy 1026 TGAATTTACATCTGCTGTTGACCAAGGTCACATTCGCTGTTGATCTTCTACTGGATA 1085
Db 481 TGAATTTACATCTGCTGTTGACCAAGGTCACATTCGCTGTTGATCTTCTACTGGATA 540
Qy 1086 TGGAGCTGATATGAATCTAGTGGCTTTGATCCACAGAGTCTAGTGGTGAAAAAGATGA 1145
Db 541 TGGAGCTGATATGAATCTAGTGGCTTTGATCCACAGAGTCTAGTGGTGAAAAAGATGA 600
Qy 1146 GCAGACATGTTGATGCGCTTATGAAAAGGGCATGATGCGCTTGTGCACACTCCTGAA 1205
Db 601 GCAGACATGTTGATGCGCTTATGAAAAGGGCATGATGCGCTTGTGCACACTCCTGAA 660
Qy 1206 GCATTTAAGAGACCAAGATGAAATGCGCTGTTGATGAAATCTCAGCCTGGAGGAGA 1265
Db 661 GCATTTAAGAGACCAAGATGAAATGCGCTGTTGATGAAATCTCAGCCTGGAGGAGA 720
Qy 1266 TGCTCTATGCTGCTGTTCCATCACCCTTGGG 1298
Db 721 AGCTCTATGCTGCTGCTCCTCACCCTTGGG 753
```

```
RESULT 3
BG209217 813 bp mRNA linear EST 21-APR-2001
LOCUS RST28732 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.
DEFINITION BG209217
ACCESSION BG209217
VERSION BG209217.1 GI:13730904
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 813)
AUTHORS Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R.,
Cain,S., Leventhal,C., Thornton,M., Ramachandran,R.,
Whittington,J., Lerner,L., Costanzo,D., McElligott,K., Booser,S.,
Mays,R., Smith,E., Veloso,N., Klika,A., Hess,J., Cothren,K., Lo,K.,
Offenbacher,J., Danzig,J. and Ducar,M.
Creation of genome-wide protein expression libraries using random
activation of gene expression
Nat. Biotechnol. 19 (5), 440-445 (2001)
21227151
Nar. Biotechnol. 19 (5), 440-445 (2001)
11329013
Contact: Scott J. Cain
Athersys, Inc.
3201 Carnegie Ave, Cleveland, OH 44115, USA
Tel: 216 431 9900
Fax: 216 361 9596
Email: scain@atersys.com
High quality sequence stop: 427.
Location/Qualifiers
1. .813
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
```

FEATURES  
source

```
ORIGIN
Query Match 27.3%; Score 682.8; DB 12; Length 813;
Best Local Similarity 95.0%; Pred. No. 3.4e-175;
Matches 780; Conservative 0; Mismatches 33; Indels 8; Gaps 7;

Qy 546 AACTGCGCTTCTTTTGAATTTGGTCTGATGATTAAGTGGTGAAGTTGGAGTAG 605
Db 1 AACTGCGCTTCTTTTGAATTTGGTCTGATGATTAAGTGGTGAAGTTGGAGTAG 60
Qy 606 ACCCTCCACCTAGCATCTGCAAAAGGATTTCTTGAATATTGCAAACTCTTTGATGGAAGA 665
Db 61 ACCCTCCACCTAGCATCTGCAAAAGGATTTCTTGAATATTGCAAACTCTTTGATGGAAGA 120
Qy 666 AGCAGCAAAAGCAGATGTAATGCTCAAGATAATGAAGACCATGTCCTCCACTCCATTCTG 725
Db 121 AGCAGCAAAAGCAGATGTAATGCTCAAGATAATGAAGACCATGTCCTCCACTCCATTCTG 180
Qy 726 TTCTCGATTGAGACCATGATATAGTTAAGTATCTGCTGCAAACTGATTTGGAAGTTCA 785
Db 181 TTCTCGATTGAGACCATGATATAGTTAAGTATCTGCTGCAAACTGATTTGGAAGTTCA 240
Qy 786 ACCTCATGTTGTTAATATCTATGGAGATACCCCTTTACACCTGGCATCTCAATGGCAA 845
Db 241 ACCTCATGTTGTTAATATCTATGGAGATACCCCTTTACACCTGGCATCTCAATGGCAA 300
Qy 846 ATTTGAAGTTGCCAAGGAATCATCCAAATATCAGAAACAGAAAGTCTGACTAAGGAAA 905
Db 301 ATTTGAAGTTGCCAAGGAATCATCCAAATATCAGAAACAGAAAGTCTGACTAAGGAAA 360
Qy 906 CATCTTCAGTGAACACAGCTTTTCATAGTGCCTTGTACCTATGSCAAGAGCATTTGACCTAGT 965
Db 361 CATCTTCAGTGAACACAGCTTTTCATAGTGCCTTGTACCTATGSCAAGAGCATTTGACCTAGT 420
Qy 966 CAAATTTCTTTGATCAGAAATGTAATAACATCAACCAAGGAGGATGGGCACAC 1025
Db 421 CAAATTTCTTTGATCAGAAATGTAATAACATCAACCAAGGAGGATGGGCACAC 480
Qy 1026 TGGATTACACTCTGCTTGTGTAACCGTCAACATTCGCTGGTTCAGTTCTTACTGGATA 1085
Db 481 TGGATTACACTCTGCTTGTGTAACCGTCAACATTCGCTGGTTCAGTTCTTACTGGATA 540
Qy 1086 TGGAGCTGATGAAATCTAGTGGCTTTGATCCAGCAGGTCTAGTGGTGAAAAAGATGA 1145
Db 541 TGGAGCTGATGAAATCTAGTGGCTTTGATCCAGCAGGTCTAGTGGTGAAAAAGATGA 599
Qy 1146 GCAGACATGTTGATGCGCTTATGAAAAGGGCATGATGCGCTTGTGCACACTCCTGAA 1205
Db 600 GCAGACTGTTGTTGATGCGCTTATGAAAAGGGCATGATGCGCTTGTGCACACTCCTG-A 658
Qy 1206 GCATTTAAGAGACCAAGATGAATTCGCTGTAATGAATTTCTCAGCTGGAGGAGA 1265
Db 659 GCATTTAAGAGACCAAGATGAATTCGCTGTAATGAATTTCTCAGCTGGAGGAGA 716
Qy 1266 TGGCTCTATGTCGTCTTCCATCACCCTTGGGAGATTAAGCATGACAAAAGAGA 1325
Db 717 AGG-TCCATATGTCGTGGTCC-TCACCTTGGGAGATTAAG--GCTGACCAAAAGAGA 772
Qy 1326 GGCAGATATTTCTCTCTCCTTAAGAGCTGGATTGCTTACATTCCATT 1366
Db 773 GGCAGATATTTCTCTCCTTAAGAGCTGGATTGCTTACATTCCATT 813
```

RESULT 4  
BG216613



Best Local Similarity 100.0%; Pred. No. 2.3e-170;  
Matches 665; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 526 TACTATGGACATGAACAGCTAACTCGCTCTTTTGAATTTGGTCTGATGAATGTA 585  
Db 1 TACTATGGACATGAACAGCTAACTCGCTCTTTTGAATTTGGTCTGATGAATGTA 60

QY 586 AGTGGTGAAGTTGGAGATAGACCCCTCCACCTAGCATCTGCAAAAGGATTTCTGAATATT 645  
Db 61 AGTGGTGAAGTTGGAGATAGACCCCTCCACCTAGCATCTGCAAAAGGATTTCTGAATATT 120

QY 646 GCAAACTCTTGATGAAGAAGGAGCAGCAAGATGGAATGCTCAAGATAAAGAGAC 705  
Db 121 GCAAACTCTTGATGAAGAAGGAGCAGCAAGATGGAATGCTCAAGATAAAGAGAC 180

QY 706 CATGTGCCACTCCATTTCTGTTCTGATTTGACACCATGATATAGTTAAGTAATCTGCTG 765  
Db 181 CATGTGCCACTCCATTTCTGTTCTGATTTGACACCATGATATAGTTAAGTAATCTGCTG 240

QY 766 CAAAGTATTTGAAGTTCAACCTCATGTTGTTAATATCTATGGAGATACCCCTTACAC 825  
Db 241 CAAAGTATTTGAAGTTCAACCTCATGTTGTTAATATCTATGGAGATACCCCTTACAC 300

QY 826 CTGGCATGCTACAAATGCAAAATTTGAAGTTGCCAAGGAATCATCAAAATATCAGGAACA 885  
Db 301 CTGGCATGCTACAAATGCAAAATTTGAAGTTGCCAAGGAATCATCAAAATATCAGGAACA 360

QY 886 GAAAGTCTGACTAAGGAAACATCTTTCACTGAAACAGCTTTTTCATAGTCTTTACCTAT 945  
Db 361 GAAAGTCTGACTAAGGAAACATCTTTCACTGAAACAGCTTTTTCATAGTCTTTACCTAT 420

QY 946 GGCAGAGCATTTCACTAGTCAAAATTTCTTCTGATCAGAAATGCTATAAATCAACACAC 1005  
Db 421 GGCAGAGCATTTCACTAGTCAAAATTTCTTCTGATCAGAAATGCTATAAATCAACACAC 480

QY 1006 CAAGGAAGGATGGGACACACTGATTACACTCTGCTTGTACACGGTCACTTCCGCTG 1065  
Db 481 CAAGGAAGGATGGGACACACTGATTACACTCTGCTTGTACACGGTCACTTCCGCTG 540

QY 1066 GTTCAGTTCTTACTGATATGAGCTGATATGAATCTAGTGGCTTGTATCCAGCAGG 1125  
Db 541 GTTCAGTTCTTACTGATATGAGCTGATATGAATCTAGTGGCTTGTATCCAGCAGG 600

QY 1126 TCTAGTGGTGAAGAAGATGAGCAGACATGTTTGAATGTTGGCTTATGAAGAAGGATCAT 1185  
Db 601 TCTAGTGGTGAAGAAGATGAGCAGACATGTTTGAATGTTGGCTTATGAAGAAGGATCAT 660

QY 1186 GCCAT 1190  
Db 661 GCCAT 665

RESULT 6  
BE897696  
LOCUS 601439226F1 NIH\_MGC\_72 Homo sapiens cDNA clone IMAGE:3924055 5',  
DEFINITION mRNA sequence.  
ACCESSION BE897696  
VERSION BE897696.1 GI:10363419  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 902)  
NIH-MGC <http://mgi.nci.nih.gov/>.  
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgabs-k@mail.nih.gov  
Tissue Procurement: ATCC/DCTD/DRP  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLN9761 row: d column: 08  
High quality sequence stop: 689.  
Location/Qualifiers  
1. .902  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3924055"  
/tissue\_type="melanotic melanoma"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH MGC\_72"  
/note="Organ: skin; Vector: pCMV-Sport6; Site: 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 2 kb. Library constructed by Life Technologies."

ORIGIN  
Query Match 24.7%; Score 618.6; DB 10; Length 902;  
Best Local Similarity 96.2%; Pred. No. 1.3e-157;  
Matches 678; Conservative 0; Mismatches 19; Indels 8; Gaps 4;

QY 332 AGGATAATGCAGAAATGATCACTTCTGCTTCACAGTGGAGCTGATATACAGAGGTTG 391  
Db 44 AGGATAATGCAGAAATGATCACTTCTGCTTCACAGTGGAGCTGATATACAGAGGTTG 103

QY 392 GATACCGTGGCTCCTCACTGCCCTCCATATTGCTTACAAATAGCTGGCCACCTAGAGGCTGCTG 451  
Db 104 GATACCGTGGCTCCTCACTGCCCTCCATATTGCTTACAAATAGCTGGCCACCTAGAGGCTGCTG 163

QY 452 ATGTGCTGTTGCAACATGGAGCTAATGCAATATTCAGAGTGCAGTCTTTTCACTCCAT 511  
Db 164 ATGTGCTGTTGCAACATGGAGCTAATGCAATATTCAGAGTGCAGTCTTTTCACTCCAT 223

QY 512 TGCAATATTCAGCGTACTATGCACATGAACAGTAACTCGCCTCTTTTGAATTTGGTG 571  
Db 224 TGCAATATTCAGCGTACTATGCACATGAACAGTAACTCGCCTCTTTTGAATTTGGTG 283

QY 572 CTGATCTAAATGTAAAGTTGAGATAGACCCCTCCACCTAGCATCTGCAAAAG 631  
Db 284 CTGATCTAAATGTAAAGTTGAGATAGACCCCTCCACCTAGCATCTGCAAAAG 343

QY 632 GATTCCTGAATATTCGAAACCTTTGATGGAAGGAGGAGCAAGAGCAGATGTGAATGCTC 691  
Db 344 GATTCCTGAATATTCGAAACCTTTGATGGAAGGAGGAGCAAGAGCAGATGTGAATGCTC 403

QY 692 AAGATAATGAAGACCATGTCCCACTCCATTTCTGTTCTCGATTTGGACACCATGATATAG 751  
Db 404 AAGATAATGAAGACCATGTCCCACTCCATTTCTGTTCTCGATTTGGACACCATGATATAG 463

QY 752 TTAAGTATCTGCTGCAAGTGAATTTGGAAGTTCAACCTCATGTTGTTAATCTATGGAG 811  
Db 464 TTAAGTATCTGCTGCAAGTGAATTTGGAAGTTCAACCTCATGTTGTTAATCTATGGAG 523

QY 812 ATACCCCTTACACCTGGCATGCTACAATGGCAAAATTTGAAGTTGCCAAGGAAATCATCC 871  
Db 524 ATACCCCTTACACCTGGCATGCTACAATGGCAAAATTTGAAGTTGCCAAGGAAATCATCC 582

QY 872 AAATATCAGAAACAGAACTCTGACTAAGAAAAAATCTTCACTAGTGAACACGCTTTTCATA 931  
Db 583 AAATATCAGAAACAGAACTCTGACTAAGAAAAA-ATCTTCACTAGTGAACACGCTTTTCATA 641

QY 932 GTGCTGTACCTATGGCAAGACATTTGACCTAGTCAAAATTTCTTCTGATCAGAAATGTCA 991  
Db 642 GTGCTGTACCTATGGCAAGACATTTGACCTAGTCAAAATTTCTTCTGATCAGAAATGTCT 698

QY 992 TAAACATCAACCAAGGAGGATGGGACACTGGATTACAT 1036  
Db 699 TAGACATTCACACACAGGAA---GGATGGCACACTGGATTACTCT 740

```

RESULT 7
BX502807      656 bp      mRNA      linear      EST 04-SEP-2003
LOCUS      DKFZp779C2070 r1 779 (synonym: hnccl) Homo sapiens cDNA clone
DEFINITION      DKFZp779C2070 5', mRNA sequence.
ACCESSION      BX502807
VERSION      BX502807.1 GI:32024614
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS      Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
TITLE      1 (bases 1 to 656)
JOURNAL      Bahr, A., Lauber, J., Mewes, H.W., Weil, B., Amid, C., Osanger, A.,
COMMENT      EST (Bahr, A., Lauber, J., Mewes, H.W., Weil, B., et al.)
CONTACT      Unpublished (2003)
CONTACT      Contact: MIPS
MIPS      Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by Qiagen (Hilden/Germany) within the cDNA sequencing
consortium of the German Genome Project.
No sl sequence available.
This clone (DKFZp779C2070) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES
source
1..656
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp779C2070"
/tissue_type="liver"
/dev_stage="fetal"
/lab_host="DH10B"
/clone_lib="779 (synonym: hnccl)"
/notes="Vector: pSport1_Sfi; Site_1: SfiI; Site_2: SfiIB"

ORIGIN
Query Match      23.4%; Score 587.4; DB 13; Length 656;
Best Local Similarity 99.7%; Pred. No. 4e-149;
Matches 588; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 ATGGGAATTATTAATCTAGACCAACCAACTGTACTGATGATGAATGGAGAAAAGTC 60
DB      67 ATGGGAATTATTAATCTAGACCAACCAACTGTACTGATGAATGGAGAAAAGTC 126
QY      61 AGTGAATCATATGTTATACAAATAGAAAGATTAGAGATGACCTGCAGATCAAGGAAAA 120
DB      127 AGTGAATCATATGTTATACAAATAGAAAGATTAGAGATGACCTGCAGATCAAGGAAAA 186
QY      121 GAATGACAGAACTAAGAAATATATTTGGCTCTGATGAAGCCTTCAGTAAGTCAATTTA 180
DB      187 GAATGACAGAACTAAGAAATATATTTGGCTCTGATGAAGCCTTCAGTAAGTCAATTTA 246
QY      181 AATTACCCGCTGAAATGGGCTGCTCTACTTCATTTATGTTGCAATTTGTGGAGCAAG 240
DB      247 AATTACCCGCTGAAATGGGCTGCTCTACTTCATTTATGTTGCAATTTGTGGAGCAAG 306
QY      241 AAATCACAATTCGAACTCTTATGTTGAAAGGGCTCGCCCATCTCGATGACAGAAAT 300
DB      307 AAATCACAATTCGAACTCTTATGTTGAAAGGGCTCGCCCATCTCGATGACAGAAAT 366
QY      301 GGATTTACAGCTTGCATTTAGCAATTTACAGGATAATGCAAGTTGATCACTTCTCTG 360
DB      367 GGATTTACAGCTTGCATTTAGCAATTTACAGGATAATGCAAGTTGATCACTTCTCTG 426
QY      361 CTTACAGTGGAGCTGATATACAGAGGTTGGATACGGTGGCTCTACCTCCCTCCATATT 420

```

```

DB      427 CTTACAGTGGAGCTGATATATACAGCAGGTTGGATACGGTGGCTCTACTCCCTCCATATT 486
QY      421 GCTACAATAGCTGCCACCTAGAGGCTGTGATGTCTCTTGCACATGGAGCTAATGTC 480
DB      487 GCTACAATAGCTGCCACCTAGAGGCTGTGATGTCTCTTGCACATGGAGCTAATGTC 546
QY      481 AATATTCAAGATGACGATTTTTCACCTCCATTCGATTTGCGAGCTGATCATGACATGAA 540
DB      547 AATATTCAAGATGACGATTTTTCACCTCCATTCGATTTGCGAGCTGATCATGACATGAG 606
QY      541 CAGGTAACCTCCCTCTTTTGAATTTGGTGTGATGATAAATGTAAGTGG 590
DB      607 CAGGTAACCTCCCTCTTTTGAATTTGGTGTGATGATAAATGTAAGTGG 656

AL599654      638 bp      mRNA      linear      EST 04-SEP-2003
DEFINITION      DKFZp313B0430 r1 313 (synonym: hlcc2) Homo sapiens cDNA clone
ACCESSION      AL599654
VERSION      AL599654
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS      Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
TITLE      1 (bases 1 to 638)
JOURNAL      Koehrer, K., Beyer, A., Mewes, H.W., Weil, B. and Wiemann, S.
COMMENT      EST (Koehrer, K., Beyer, A., Mewes, H.W., Weil, B. and Wiemann, S.)
CONTACT      Unpublished (1999)
CONTACT      Contact: MIPS
MIPS      Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by BMFZ (Biomedical Research Center at the Charite,
Berlin/Germany) within the cDNA sequencing consortium of the German
Genome Project.
No sl sequence available.
This clone (DKFZp313B0430) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES
source
1..638
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp313B0430"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="313 (synonym: hlcc2)"
/notes="Vector: pTriplex2; Site_1: SfiI; Site_2: SfiIB;
cDNA-collection"

ORIGIN
Query Match      22.2%; Score 557; DB 9; Length 638;
Best Local Similarity 100.0%; Pred. No. 8.3e-141;
Matches 557; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ATGGGAATTATTAATCTAGACCAACCAACTGTACTGATGATGAAGAAAAGTC 60
DB      82 ATGGGAATTATTAATCTAGACCAACCAACTGTACTGATGATGAAGAAAAGTC 141
QY      61 AGTGAATCATATGTTATCAATAGAAAGATTAGAGATGACCTGCAGATCAAGGAAAA 120
DB      142 AGTGAATCATATGTTATCAATAGAAAGATTAGAGATGACCTGCAGATCAAGGAAAA 201
QY      121 GAATGACAGAACTAAGAAATATATTTGGCTCTGATGAGCTTCAGTAAAGTCAATTTA 180
DB      202 GAATGACAGAACTAAGAAATATATTTGGCTCTGATGAGCTTCAGTAAAGTCAATTTA 261

```

QY 181 AATTACCGCAGTGAATGGCTGTCTCTACTTCAATTTATGTCATTTGTCATTTGGGAGCAAG 240  
Db 262 AATTACCGCAGTGAATGGCTGTCTCTACTTCAATTTATGTCATTTGTCATTTGGGAGCAAG 321  
QY 241 AATACATATTCGAATCTTATGTTGAAGAGGCTCCGCCCATCTCGACTGACAGAAAT 300  
Db 322 AATACATATTCGAATCTTATGTTGAAGAGGCTCCGCCCATCTCGACTGACAGAAAT 381  
QY 301 GGNATTACAGCCTTGCATTTACAGCTTTACAGGATATGAGGATTCGATCACTTCTCTG 360  
Db 382 GGNATTACAGCCTTGCATTTACAGCTTTACAGGATATGAGGATTCGATCACTTCTCTG 441  
QY 361 CTTACAGTGGAGCTGATATACAGAGGTTGGATACGGTGGCTCCTCAGTCCCTCCATATT 420  
Db 442 CTTACAGTGGAGCTGATATACAGAGGTTGGATACGGTGGCTCCTCAGTCCCTCCATATT 501  
QY 421 GTCACATAGTGGCCACCTAGAGGCTGCTGATGCTGTTGTCACATGAGAGCTAATGTC 480  
Db 502 GTCACATAGTGGCCACCTAGAGGCTGCTGATGCTGTTGTCACATGAGAGCTAATGTC 561  
QY 481 AATATTCAAGATGCGAGTCTTTTTCATCCTCATTTGATGAGGCTTATGAGCATGAA 540  
Db 562 AATATTCAAGATGCGAGTCTTTTTCATCCTCATTTGATGAGGCTTATGAGCATGAA 621  
QY 541 CAGGTAACCTGGCTTCT 557  
Db 622 CAGGTAACCTGGCTTCT 638

RESULT 9  
BX112488  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

112488  
Soares fetal lung NbHL19W Homo sapiens cDNA clone  
IMAGE:1931107, mRNA sequence.  
112488.1 GI:27878929  
EST.  
Homo sapiens (human)  
Homo sapiens

REFERENCE  
AUTHORS  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 491)  
Radelof, U., Schneider, D. and Korn, B.  
Human Unigeneset - R2PD3  
Unpublished (2003)  
Contact: Ina Rols  
R2PD Deutsches Ressourcenzentrum fuer Genomforschung GmbH  
Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany  
R2PD; IMAGE:1931107, mRNA sequence.  
R2PD; IMAGE:1931107, mRNA sequence.  
Human Unigeneset - R2PD3 (R2PD; No. 972)  
http://www.rzpd.de/CloneCards/cgi-bin/showLib.pl.cgi/response?libNo=972 Contact: Ina Rols  
R2PD Deutsches Ressourcenzentrum fuer Genomforschung GmbH  
Heubnerweg 6, D-14059 Berlin, Germany  
Tel: +49 30 32639 101  
Fax: +49 30 32639 111  
www.rzpd.de

FEATURES  
source  
1. 491  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:1931107"  
/dev\_stage="19 weeks"  
/lab\_host="DEH10B (ampicillin resistant)"  
/clone\_lib="Soares\_fetal\_lung\_NbHL19W"  
/notes="Organ: lung; Vector: pT73D (Pharmacia) with a  
modified polylinker; Site\_1: Not I; Site\_2: Eco RI; 1st

strand cDNA was primed with a Not I - oligo (dT) primer  
[5'-TGTTACCAATCTGAAGTGGAGCGCGCAATTTTCTTTT-3'],  
double-stranded cDNA was size selected, ligated to Eco RI  
adapters (Pharmacia), digested with Not I and cloned into  
the Not I and Eco RI sites of a modified pT73 vector  
(Pharmacia). Library went through one round of  
normalization to a Cot = 5. Library constructed by Bento  
Soares and M. Fatima Bonaldo. This library was constructed  
from the same fetus as the fetal heart library, Soares  
fetal heart NbHL19W."

ORIGIN

Query Match 19.6%; Score 491; DB 13; Length 491;  
Best Local Similarity 100.0%; Pred. No. 8.5e-123;  
Matches 491; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1109 CTTGTGATCCAGCAGCTCTAGTGGTGAAGAAGATGAGCAGACATGTTGATGGGCTT 1168  
Db 1 CTTGTGATCCAGCAGCTCTAGTGGTGAAGAAGATGAGCAGACATGTTGATGGGCTT 60  
QY 1169 ATGAAAAAGGGCATGATGCCATTGTGCACCTCCTGAGCATTATAAGAGAGCCACAAGATG 1228  
Db 61 ATGAAAAAGGGCATGATGCCATTGTGCACCTCCTGAGCATTATAAGAGAGCCACAAGATG 120  
QY 1229 AATTGCCCTGTAATGAATATTTCTAGCCTGGAGAGATGGCTCTCTATGTCGTCTTCAT 1288  
Db 121 AATTGCCCTGTAATGAATATTTCTAGCCTGGAGAGATGGCTCTCTATGTCGTCTTCAT 180  
QY 1289 CACCTTTGGGGAAGATTAAAGCATGACAAAGAGAGAGCAGATATTCCTCTCTAAGAG 1348  
Db 181 CACCTTTGGGGAAGATTAAAGCATGACAAAGAGAGAGCAGATATTCCTCTCTAAGAG 240  
QY 1349 CTGGATTGCTTCAATTTCCATTTCCATCTCAGCTCTCAGAAATTTAGTTCATGAGATTATG 1408  
Db 241 CTGGATTGCTTCAATTTCCATTTCCATCTCAGCTCTCAGAAATTTAGTTCATGAGATTATG 300  
QY 1409 GCTCAGGTTCTTTTGGGAAGTATATAAGAGCAGATGAGAAATAAATAGTGGCTATAA 1468  
Db 301 GCTCAGGTTCTTTTGGGAAGTATATAAGAGCAGATGAGAAATAAATAGTGGCTATAA 360  
QY 1469 AAGTTATCGAGCCAAATACCTACTCTCTCAAGTCAGATGTGGATATGTTTCCGAGAGG 1528  
Db 361 AAGTTATCGAGCCAAATACCTACTCTCTCAAGTCAGATGTGGATATGTTTCCGAGAGG 420  
QY 1529 TGTCCATTTCTGCGAGCTCAATCATCTCTGCGTAATTTAGTTCGCTGCTTGA 1588  
Db 421 TGTCCATTTCTGCGAGCTCAATCATCTCTGCGTAATTTAGTTCGCTGCTTGA 480  
QY 1589 ATGATCCCAGC 1599  
Db 481 ATGATCCCAGC 491

RESULT 10  
BU622977/c  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

622977  
UI-H-FL1-bgb-n-08-0-UI-s1 NCI CGAP FL1 Homo sapiens cDNA clone  
UI-H-FL1-bgb-n-08-0-UI 3', mRNA sequence.  
BU622977  
EST.  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 712)  
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)  
Contact: Robert Strausberg, Ph.D.  
Email: cgaps-r@mail.nih.gov  
Tissue Procurement: James Martin

712 bp mRNA linear EST 23-SEP-2002  
UI-H-FL1-bgb-n-08-0-UI-s1 NCI CGAP FL1 Homo sapiens cDNA clone  
UI-H-FL1-bgb-n-08-0-UI 3', mRNA sequence.

BU622977  
UI-H-FL1-bgb-n-08-0-UI-s1 NCI CGAP FL1 Homo sapiens cDNA clone  
UI-H-FL1-bgb-n-08-0-UI 3', mRNA sequence.  
BU622977  
EST.  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 712)  
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)  
Contact: Robert Strausberg, Ph.D.  
Email: cgaps-r@mail.nih.gov  
Tissue Procurement: James Martin

cDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: Clone distribution information can be obtained  
 from Dr. M. Bento Soares, bento-soares@uiowa.edu  
 The following repetitive elements were found in this cDNA  
 sequence: 17-303, >LTR8#LTR/Retroviral (matched complement)  
 Seq primer: M13 FORWARD  
 POLYA=Yes.

## FEATURES

source

Location/Qualifiers

1..712  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clones="UI-H-FL1-bgp-n-08-0-UI"  
 /tissue\_type="Cell lines"  
 /dev\_stage="adult"  
 /lab\_host="DH10B (Life Technologies)"  
 /clone\_lib="NCI\_CGAP\_FLI"  
 /note="Organ: Chondrosarcoma; Vector: pT73-Pac  
 (pharmacia) with a modified polylinker; Site 1: EcoR I;  
 Site 2: Not I; NCI CGAP FL1 is a normalized cDNA library  
 derived from a pool of mRNA obtained from 4 cell lines  
 from grade III chondrosarcoma tissues. The library was  
 constructed according to Bonaldo, Lennon and Soares,  
 Genome Research, 6:791-806, 1996. First strand cDNA  
 synthesis was primed with an oligo-dT primer containing a  
 Not I site. Double stranded cDNA was ligated to an EcoR I  
 adaptor, digested with Not I, and cloned directionally  
 into pT73-Pac vector. The oligonucleotide used to prime  
 the synthesis of first-strand cDNA contains a library tag  
 sequence that is located between the Not I site and the  
 (d)18 tail. The sequence tag for this library is  
 GAGGTCGGTG. The cell lines were provided by Dr. James  
 Martin from the University of Iowa.  
 TAG TISSUE=Human Chondrosarcoma Grade 3 cell line mix  
 TAG LIB=UI-H-FL1  
 TAG\_SEQ=GAGGTCGGTG"

## ORIGIN

Query Match 16.4%; Score 410.4; DB 13; Length 712;  
 Best Local Similarity 99.8%; Pred. No. 1.2e-100;  
 Matches 411; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1361 CACATTTCCATCTTCAGCTCTCAGAAATTCAGTTCCTCAGATGATTATGGCTCAGGTTCTT 1420  
 DB 712 CACATTTCCATCTTCAGCTCTCAGAAATTCAGTTCCTCAGATGATTATGGCTCAGGTTCTT 653

QY 1421 TTGGGAAAGTATATAAAGGACGATGCGAGAAATAAATAGTGGCTATPAAACGTTATCGAG 1480  
 DB 652 TTGCGAAGTATATAAAGGACGATGCGAGAAATAAATAGTGGCTATPAAACGTTATCGAG 593

QY 1481 CCATACCTACTGCTCCAACTCAGATGCGATGCTATGTTTCCGAGAGGTGTCATCTCT 1540  
 DB 592 CCAATACCTACTGCTCCAACTCAGATGCGATGCTATGTTTCCGAGAGGTGTCATCTCT 533

QY 1541 GCCAGCTCAATCATCTCCCTGGGTAAATTCAGTTCCTGGGTGCTTGGTGAATGATCCAGCC 1600  
 DB 532 GCCAGCTCAATCATCTCCCTGGGTAAATTCAGTTCCTGGGTGCTTGGTGAATGATCCAGCC 473

QY 1601 AGTTGGCATTGTCACCTCAATATACATATCAGGGGGTCTCTGTTCCTCCCTCTTATGAGC 1660  
 DB 472 AGTTGGCATTGTCACCTCAATATACATATCAGGGGGTCTCTGTTCCTCCCTCTTATGAGC 413

QY 1661 AGAAGAGATTCTTGATTGTCAGTCTAAATTAATTATTCAGTAGATGTTGCCAAGGCA 1720  
 DB 412 AGAAGAGATTCTTGATTGTCAGTCTAAATTAATTATTCAGTAGATGTTGCCAAGGCA 353

QY 1721 TGGAGTACCTTTCACAACTGACACAGCCCAATTATACATCGTGACTTGAACAG 1772  
 DB 352 TGGAGTACCTTTCACAACTGACACAGCCCAATTATACATCGTGACTTGAACAG 301

## RESULT 11

BF556094

LOCUS

DEFINITION UI-R-Al-eb-g-10-0-UI r1 UI-R-Al Rattus norvegicus cDNA clone  
 EST 12-DEC-2000

ACCESSION

BF556094

VERSION

BF556094.1

KEYWORDS

SOURCE

ORGANISM

Rattus norvegicus (Norway rat)

REFERENCE

1 (bases 1 to 507)

AUTHORS

Bonaldo,M.F., Lennon,G. and Soares,M.B.

TITLE

Normalization and subtraction: two approaches to facilitate gene

discovery

JOURNAL

Genome Res. 6 (9), 791-806 (1996)

MEDLINE

97044477

PUBMED

COMMENT

Contact: Soares, MB

Coordinated Laboratory for Computational Genomics

University of Iowa

375 Newton Road , 4156 MEBRF, Iowa City, IA 52242, USA

Tel: 319 335 8250

Fax: 319 335 9565

Email: bento-soares@uiowa.edu

cDNA Library Preparation: M.B. Soares Lab Clone distribution:

clones will be available through Research Genetics (www.resgen.com)

This clone is also available through the I.M.A.G.E. Consortium at

LNL (info@image.llnl.gov). IMAGE ID= 1771245

Seq primer: M13 Forward

Location/Qualifiers

1..507

/organism="Rattus norvegicus"

/mol\_type="mRNA"

/strain="Sprague-Dawley"

/db\_xref="taxon:10116"

/clone="UI-R-Al-eb-g-10-0-UI"

/dev\_stage="adult"

/lab\_host="DH10B (Life Technologies)"

/clone\_lib="UI-R-Al"

/note="Vector: pT73D-Pac (Pharmacia) with a modified

polylinker; Site 1: Not I; Site 2: Eco RI; The UI-R-Al

library is a subtracted library derived from the UI-R-A0

library. The UI-R-A0 library consisted of a mixture of

individually tagged normalized libraries constructed from

rat placenta, adult lung, brain, liver, kidney, heart,

spleen, ovary, and muscle. The tag is a string of 3-5

nucleotides present between the Not I site and the

oligo-dT track which allows identification of the library

of origin of a clone within the mixture. The subtracted

library (UI-R-Al) was constructed as follows: PCR

amplified cDNA inserts from a pool of approximately 3,840

UI-R-A0 clones from which 3, ESTs had been derived was

used as a driver in a hybridization with the UI-R-A0

library in the form of single-stranded circles. The

remaining single-stranded circles (subtracted library) was

purified by hydroxyapatite column chromatography,

converted to double-stranded circles and electroporated

into DH10B bacteria (Life Technologies) to generate the

UI-R-Al library. This procedure has been previously

described (Bonaldo, Lennon and Soares, Genome Research 6:

791-806, 1996)"

## ORIGIN

Query Match 15.7%; Score 392.8; DB 10; Length 507;  
 Best Local Similarity 86.6%; Pred. No. 6.5e-96;  
 Matches 433; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

QY 1702 GTAGATGTTGCCAAGGCGATGGAGTACCTTCACACCTTGACAGCCCAATTATACATCGT 1761  
 DB 8 GTAGATGTTGCCAAGGCGATGGAGTACCTTCACACCTTGACAGCCCAATTATACATCGC 67

1762 GACTGAACAGTCACAAATTTCTTCTATAGGATGGGCATGCTGTGGTGGAGATTTT 1821  
Db |||||  
68 GACCTGAACAGGCACAAATTTCTGCTCTATAGGATGGCCATGCTGTGGTGGAGATTTT 127  
Qy |||||  
1822 GGAGATCAAGATTTCTACAGTCTCTGGATGAAGACACATGACAAACACCTGGGAAC 1881  
Db |||||  
128 GGAGATCAAGATTTCTGAGTCCCTGGATGAAGACACATGACAAACACCTGGGAAC 187  
Qy |||||  
1882 CTCGGTGGATGGCTCTCTGAGTGTTCACGCGAGTGCATCGGTACACCATCAAGAGCAGAT 1941  
Db |||||  
188 CTCGGTGGATGGCCCTCAGGTGTTCACACAGTGCACAGATACACCATCAAGGCTGAT 247  
Qy |||||  
1942 GCTTTCAGTATGCTGTGTGTGGGAAATTCATCTGGGAAATTCGATTCGCTCAT 2001  
Db |||||  
248 GCTTTCAGTATGCTGTGTGTGGGAGTCCCTCACTGGAGAAATTCGATTCGCTCAT 307  
Qy |||||  
2002 CTCAGCCAGCGGCTCGGCGACGACATGCGTTCACCCACATCAGACCTCCCATTTGGC 2061  
Db |||||  
308 CTCAGCCAGCGGCTCGGCGACGACATGCGTTCACCCACATCAGACCTCCCATCGGC 367  
Qy |||||  
2062 TATTCATTTCCAGCCCATATCATCTCTGTGTGATGAGGGTGAAGCGATGCTCGAA 2121  
Db |||||  
368 TATTCATTTCCAGCCCATATCATCTCTGTGTGATGAGGGTGAAGCGATGCTCGAA 427  
Qy |||||  
2122 GGAGACCCGAAATTTCTGAAGTGTCTGAAGTTAGAGAGTGTCTCTGCAACATTGAG 2181  
Db |||||  
428 GGAGACCCGAAATTTCTGAAGTGTCTGAAGTGTAGAGAGTGTCTCTGCAACATTGAG 487  
Qy |||||  
2182 CTGATGCTCTCGCATCAAG 2201  
Db |||||  
488 CTCATGCTCCAGCATCAAG 507

RESULT 12  
LOCUS AU130402 862 bp mRNA linear EST 01-AUG-2002  
DEFINITION AU130402 NT2RP3 Homo sapiens cDNA clone NT2RP3000778 5', mRNA  
sequence.

ACCESSION AU130402  
VERSION AU130402.1 GI:10990756  
KEYWORDS EST.  
SOURCE Homo sapiens  
ORGANISM Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 862)  
AUTHORS Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,  
Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and  
Ishigai,T.

TITLE HRI human cDNA project  
JOURNAL Unpublished (2000)  
COMMENT Contact: Takao Isoqai  
Genomics Laboratory  
Helix Research Institute  
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan  
Tel: 81-438-52-3975  
Fax: 81-438-52-3986  
Email: genomics@hri.co.jp  
HRI human cDNA project; 5'- & 3'-end one pass sequencing; Helix  
Research Institute; cDNA library construction; Department of  
Virology, Institute of Medical Science, University of Tokyo, and  
Helix Research Institute.

FEATURES  
source  
1. .862  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="NT2RP3000778"  
/cell\_type="teratocarcinoma"  
/cell\_line="NT2"  
/clone\_lib="NT2RP3"  
/note="Vector: pME18SFL3; mRNA from NT2 neuronal precursor

ORIGIN  
Query Match 15.3%; Score 382.4; DB 9; Length 862;  
Best Local Similarity 99.0%; Pred. No. 5.8e-93;  
Matches 394; Conservative 0; Mismatches 3; Indels 1; Gaps 1;  
cells after 2-weeks retinoic acid (RA) induction"  
Qy 38 CTGATGAATGGAGAAAAAGTCAGTGAATCATATGTTATCACAATAGAAAGATTGAAG 97  
Db |||||  
376 CTGATGAATGGAGAAAAAGTCAGTGAATCATATGTTATCACAATAGAAAGATTGAAG 435  
Qy 98 ATGACCTGCGAGATCAAGAAAAAGAACTGCAGAGAACTAAAGGAATATATTTGGCTCTCATG 157  
Db |||||  
436 ATGACCTGCGAGATCAAGAAAAAGAACTGCAGAGAACTAAAGGAATATATTTGGCTCTCATG 495  
Qy 158 AAGCCTTCAGTAAATCAATTTAAATTTACCGCAGTGAATAATGGGCTGTCTTACTTCAAT 217  
Db |||||  
496 AAGCCTTCAGTAAATCAATTTAAATTTACCGCAGTGAATAATGGGCTGTCTTACTTCAAT 555  
Qy 218 TATGTTGCATTTGCGAGCAAGAAATCACATATTCGAACTCTTATGTTGAAAGGGCTCC 277  
Db |||||  
556 TATGTTGCATTTGCGAGCAAGAAATCACATATTCGAACTCTTATGTTGAAAGGGCTCC 615  
Qy 278 GCCATCTCGATGCAAGAAATGGAATTTACAGCTTCGATTTAGCAGTTTACAGGATA 337  
Db |||||  
616 GCCATCTCGATGCAAGAAATGGAATTTACAGCTTCGATTTAGCAGTTTACAGGATA 675  
Qy 338 ATGCGAATTTGATCACTTCTCTGCTTCACAGTGGAGCTGATACAGCAGTTTGGATACG 397  
Db |||||  
676 ATGCGAATTTGATCACTTCTCTGCTTCACAGTGGAGCTGATACAGCAGTTTGGATACN 735  
Qy 398 GTGGCCTCACTGCCCTCCATATTCCTACAATAGCTGGC 435  
Db |||||  
736 GTGGCCTCACTG-CCTTCATATTCCTACAATAGCTGGC 772

RESULT 13  
LOCUS CF412400 950 bp mRNA linear EST 02-SEP-2003  
DEFINITION CH3#079 E08MF Canine heart normalized cDNA Library in paluascript  
Canis familiaris cDNA clone CH3#079\_E08 5', mRNA sequence.

ACCESSION CF412400  
VERSION CF412400.1 GI:34413646  
KEYWORDS EST.  
SOURCE Canis familiaris (dog)  
ORGANISM Canis familiaris  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
REFERENCE 1 (bases 1 to 950)  
AUTHORS Yi,Y., Desai,R., Olarte,M., Henthorn,P. and George A.L.  
TITLE Expressed sequence tags from Canine heart  
JOURNAL Unpublished (2003)  
COMMENT Other ESTs: CH3#079\_E08MR  
Contact: George Al.  
Division of Genetic Medicine  
Vanderbilt University  
529 Light Hall, 2215 Garland Avenue, Nashville, TN 37232-0275, USA  
Tel: 615 936 2660  
Fax: 615 936 2661  
Email: al.george@vanderbilt.edu  
Insert Length: 1883 Std Error: 0.00  
Seq primer: MF: GTTTTCCAGTCACGACGTTG  
High quality sequence start: 68  
High quality sequence stop: 441.  
Location/Qualifiers

FEATURES  
source  
1. .950  
/organism="Canis familiaris"  
/mol\_type="mRNA"  
/db\_xref="taxon:9615"  
/clone="CH3#079\_E08"  
/tissue\_type="heart"  
/cell\_type="heart"  
/dev\_stage="mixed developmental stages (adult, 30 day - 40

```

day fetal)"/
/clone_lib="Canine heart normalized cDNA Library in
pBluescript"
/notes="Organ: heart; Vector: pBluescript; Site 1: 5' of
vector NotI; Site 2: 3' of vector EcoRI; Tissue source:
dog heart (adult_30 day - 40 day fetal), right and left
atria and ventricle. Dog breed - mixed (beagle, German
shepherd, pointer, Irish setter). Library construction:
oligo-dr primed"

ORIGIN
Query Match 14.9%; Score 374; DB 14; Length 950;
Best Local Similarity 93.9%; Pred. No. 1.2e-90;
Matches 400; Conservative 0; Mismatches 25; Indels 1; Gaps 1;

QY 1513 ATGTTTTCGAGAGGTGTCATCTCTGCCAGCTCAATCATCCCTGCGTAATTCAGTTT 1572
Db AGGTTCTCGGAGAGGTGTCATCTCTGCCGCTCAATCATCCCTGTAATTCAGTTT 125

QY 1573 GTGGGTGCTTCTGAATGATCCAGCCAGCTTGGCCATTTGCTACTCAATACATATCAGG 1632
Db |||||

QY 126 GTGGGTGCTTCTGAATGATCCAGCCAGCTTGGCCATTTGCTACTCAATACATATCAGG 185
Db |||||

QY 1633 GGTCTCTGTTCTCCCTCCCTTCATGAGCAGAGAGGATTTCTGATTTGCGAGCTAAATTA 1692
Db GTTTCCCTGTTCTCCCTCCCTTCATGAGCAGAGAGGATTTCTGATTTGCGAGCTAAATTA 245

QY 1693 ATTATTCAGTAGATGTTGCCAAGGATGAGTACTTCAACACCTGACACAGCAATT 1752
Db |||||

QY 246 ATTATTCAGTAGAGTGTGCCAAGGATGAGTACTTCAACACCTGACACAGCAATT 305
Db |||||

QY 1753 ATACATCTGCTTGAACAGTCACATATTTCTTCTATGAGATGGCATGCTGTGGTG 1812
Db |||||

QY 306 ATACACCTGATTTAAACAGTCAATATTTCTTCTGTGAGATGGCATGCTGTGGTG 365
Db |||||

QY 1813 GCAGATTTTGAGAAATCAAGATTTCTACGTCTCTGGATGAGACAAACATGACAAACAA 1872
Db GCAGATTTTGAGAAATCAAGATTTCTACGTCTCTGGATGAGACAAACATGACAAACAA 425

QY 1873 CCTGGACCTCCGTTGGATGGCTCTGAGTGTCTCAGCGAGTGCAC-TGGGTACACCAT 1931
Db CCTGGACCTCCGTTGGATGGCTCTCAGCGAGTGTCTCAGCGAGTGTCTCAGCGAGTGTCTACACCATGACACCGGTACACCAT 485

QY 1932 CAAGC 1937
Db 486 CTAAGC 491

RESULT 14
BB658710
LOCUS
DEFINITION BB658710 RIKEN full-length enriched, 13 days embryo heart Mus
musculus cDNA clone D33006E15 5', mRNA sequence.
ACCESSION BB658710
VERSION BB658710.1 GI:16492535
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 634)
Hara, A., Hiramoto, K., Hori, F., Ishii, Y., Ito, M., Kawai, J.,
Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T.,
Konno, H., Kouda, M., Koya, S., Matsuyama, T., Miyazaki, A., Nomura, K.,
Ono, M., Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F.,
Takeda, Y., Tanaka, T., Taya, T., Muramatsu, M. and Hayashizaki, Y.
RIKEN Mouse ESTs (Arakawa, T., et al. 2001)
Unpublished (2001)
TITLE
JOURNAL
COMMENT Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute

```

```

The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsr.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, F., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
wagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,
Watabiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T.,
Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A.
and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P.,
Sugahara, Y. and Hayashizaki, Y.
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Kondo, S., Shinagawa, A., Saito, T., Kiyosawa, H., Yamanaka, I.,
Aizawa, K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K. and
Hayashizaki, Y.
Computational Analysis of Full-length Mouse cDNAs Compared with
Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)
Please visit our web site (http://genome.gsc.riken.go.jp) for
further details.
e mouse tissues.
Location/Qualifiers
1. .634
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="D33006E15"
/tissue_type="heart"
/dev_stage="13 days embryo"
/lab_host="DH10B"
/clone_lib="RIKEN full-length enriched; 13 days embryo
heart"
/notes="Site 1: SalI; Site 2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN. Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5'
GAGAGAGAGCGCGCAACTCGAGTTTCTTTTCTTTTCTTTT 3'], cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. Second strand cDNA was prepared with the
primer adapter of sequence [5'
GAGAGAGAGTTCTCGAGTTAATTAATTAATTCCTCCCTCCCTCC 3']. cDNA
was cleaved with BamHI and XhoI. Vector: a modified
pBluescript KS(+) after bulk excision from Lambda FLC I."

```

## ORIGIN

```

Query Match 14.3%; Score 358.6; DB 10; Length 634;
Best Local Similarity 79.5%; Pred. No. 1.7e-86;
Matches 424; Conservative 0; Mismatches 109; Indels 0; Gaps 0;

QY 1 ATGGGAAATTAATTAATCTAGACCAACCTTGTACTGTATGATGAGAGAAAGATC 60
Db 102 ATGGGAAATTAATTAATCTAGACCAACCTTGTACTGTATGATGAGAGAAAGATC 161

QY 61 AGTGAATCATATCTTATCAATAGAGATAGAGATGACCTGCAGATCAAGAGAAA 120
Db 162 AGTGAATCATATCTTATCAATAGAGATAGAGATGACCTGCAGATCAAGAGAAA 221

QY 121 GAATCAGAGAACTAAGGAATATATTGCTCTGATGAAGCTTCAGTAAAGTCAATTTA 180

```

```

FEATURES
source
Location/Qualifiers
1. .465
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAG9998B075155 ; IMAGE:2091174"
/lab_host="DH10B"
/clone_lib="Scorates NFL T CBC S1"
/note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with
a modified polylinker; Site_1: Not I; Site_2: Eco RI;
Equal amounts of plasmid DNA from three normalized
libraries (fetal lung NBHL19W, testis NHT, and B-cell
NCI CGAP GCBI) were mixed, and ss circles were made in

```

Search completed: September 6, 2004, 07:38:05  
Job time : 4175.17 secs

Search completed: September 6, 2004, 07:38:05  
Job time : 4175.17 secs

This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 2, 2004, 14:15:41 ; Search time 136 Seconds  
(without alignments)  
1937.189 Million cell updates/sec

Title: US-10-626-173-2

Perfect score: 4390

Sequence: 1 MGNYKSRPTQTCDEWKKV.....PMSSMHFHSRNSSEFEDSS 835

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_25.\*

- 1: sp\_archaea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mhc.\*
- 8: sp\_organelle.\*
- 9: sp\_phase.\*
- 10: sp\_plant.\*
- 11: sp\_rodent.\*
- 12: sp\_virus.\*
- 13: sp\_vertebrate.\*
- 14: sp\_unclassified.\*
- 15: sp\_virus.\*
- 16: sp\_bacteriap.\*
- 17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4390	100.0	835	4 Q9Y2V6	Q9Y2V6 homo sapien
2	4079	92.9	835	11 Q7TQP6	Q7TQP6 rattus norv
3	1716.5	39.1	950	5 O17055	O17055 caenorhabdi
4	672	15.3	165	11 O8BFR0	O8BFR0 mus muscucu
5	486	11.1	546	10 O22558	O22558 arabidopsis
6	482	11.0	603	10 Q7XXN2	Q7XXN2 oryza sativ
7	476.5	10.9	445	10 Q9LMF8	Q9LMF8 arabidopsis
8	476	10.8	525	10 Q9FIL6	Q9FIL6 arabidopsis
9	475.5	10.8	470	10 Q7X153	Q7X153 oryza sativ
10	475	10.8	438	10 Q9X187	Q9X187 arabidopsis
11	475	10.8	438	10 Q89J9	Q89J9 arabidopsis
12	473.5	10.8	553	10 O81808	O81808 arabidopsis
13	472	10.8	525	10 Q8W022	Q8W022 arabidopsis
14	469.5	10.7	570	10 Q8RWL6	Q8RWL6 arabidopsis
15	469.5	10.7	638	5 Q86AT8	Q86AT8 dictyostell
16	460.5	10.5	1719	4 Q13768	Q13768 homo sapien

17	454	10.3	1856	4	Q99407	Q99407 homo sapien
18	446	10.2	1549	5	Q9V4B1	Q9V4B1 drosophila
19	444	10.1	1136	6	Q9N180	Q9N180 bos taurus
20	444	10.1	1848	11	Q61302	Q61302 mus muscucu
21	443	10.1	1549	5	Q24241	Q24241 drosophila
22	442	10.1	422	10	Q8GV29	Q8GV29 oryza sativ
23	442	10.1	422	10	O7XPE4	O7XPE4 oryza sativ
24	442	10.1	743	5	Q9Y1Y2	Q9Y1Y2 ephydatia f
25	439.5	10.0	462	10	Q39886	Q39886 glycine max
26	438.5	10.0	417	10	Q8GV30	Q8GV30 oryza sativ
27	438.5	10.0	616	13	O73613	O73613 xenopus lae
28	438	10.0	412	10	Q9M085	Q9M085 arabidopsis
29	436.5	9.9	919	4	Q8N8A2	Q8N8A2 homo sapien
30	436	9.9	411	10	O9ZQ31	O9ZQ31 arabidopsis
31	435.5	9.9	638	4	Q81Z72	Q81Z72 homo sapien
32	435	9.9	1863	4	Q7Z3L5	Q7Z3L5 homo sapien
33	434.5	9.9	541	10	Q7XSP3	Q7XSP3 oryza sativ
34	431.5	9.8	1943	11	Q61307	Q61307 mus muscucu
35	430	9.8	1219	11	Q8C8R3	Q8C8R3 mus muscucu
36	428.5	9.8	1726	11	O8VC68	O8VC68 mus muscucu
37	428.5	9.8	2622	11	O70511	O70511 rattus norv
38	427.5	9.7	416	10	Q94C42	Q94C42 triticum ae
39	427.5	9.7	1088	4	Q13484	Q13484 homo sapien
40	426	9.7	448	5	Q9V400	Q9V400 drosophila
41	426	9.7	961	5	Q86B19	Q86B19 dictyostell
42	423.5	9.6	615	13	Q801N1	Q801N1 xenopus lae
43	423.5	9.6	1897	4	Q723G4	Q723G4 homo sapien
44	423.5	9.6	1762	11	O88521	O88521 rattus norv
45	420.5	9.6	1009	5	Q8SWY2	Q8SWY2 drosophila

#### ALIGNMENTS

#### RESULT 1

Q9Y2V6 PRELIMINARY; PRT; 835 AA.

AC Q9Y2V6;  
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Hypothetical protein (Cardiac ankyrin repeat kinase).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Heart;  
 RA Wei Y.J., Ding J.F., Xiong H., Zhou Y., Hui R.T., Liew C.C.;  
 RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Jeyaseelan R.;  
 RT "Cardiac Ankyrin Repeat Kinase (CARK)";  
 RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF116826; AAD29632.1; -;  
 DR EMBL; AY303691; AAF72030.1; -;  
 DR HSSP; Q00420; IAWC.  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.  
 DR GO; GO:0016740; F:transferase activity; IEA.  
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.  
 DR InterPro; IPR002110; ANK.  
 DR InterPro; IPR000719; Prot\_kinase.  
 DR InterPro; IPR01245; Tyr\_kinase.  
 DR Pfam; PF00023; ank; 10.  
 DR Pfam; PF00069; pkinase; 1.  
 DR PRINTS; PR01415; ANKYIN.  
 DR PRINTS; PR0109; TYRKINASE.  
 DR ProDom; PDD000001; Prot\_kinase.  
 DR SMART; SMO0248; ANK; 8.  
 DR PROSITE; PS50085; ANK\_REPEAT; 6.

DR PROSITE; P500297; ANK REP REGION; 1.  
DR PROSITE; P500107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; P500111; PROTEIN\_KINASE\_DOM; 1.  
KW Hypothetical protein; ANK repeat; ATP-binding; Kinase; Repeat;  
KW Transferrase; Tyrosine-protein kinase.  
SQ SEQUENCE 835 AA; 92850 MW; 3B21484B434F46B8 CRC64;

Query Match 100.0%; Score 4390; DB 4; Length 835;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 835; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGNYSRPTQTCTDEWKKKVSSEYVITIERLEDLQIKKELTELNRNIFGSDFAFSKYNL 60  
Db |||||  
QY 1 MGNYSRPTQTCTDEWKKKVSSEYVITIERLEDLQIKKELTELNRNIFGSDFAFSKYNL 60  
Db |||||

QY 61 NYRTENGSLHLCCICCGKKSHIRTLMLKGLRSPRLTRNGFTALHLAVYKDAEITSL 120  
Db |||||  
QY 61 NYRTENGSLHLCCICCGKKSHIRTLMLKGLRSPRLTRNGFTALHLAVYKDAEITSL 120  
Db |||||

QY 121 LHSAGADIQQVGYGGLTALHIAITAGHLEAADVLLQHGANNVNIQDAVFTPLHIAAYGHE 180  
Db |||||  
QY 121 LHSAGADIQQVGYGGLTALHIAITAGHLEAADVLLQHGANNVNIQDAVFTPLHIAAYGHE 180  
Db |||||

QY 181 QVTRLLKFGADVNSGEVGRPLHLASAKGFNIAKLLMEBGSKADVNAQDNEDHVPFLH 240  
Db |||||  
QY 181 QVTRLLKFGADVNSGEVGRPLHLASAKGFNIAKLLMEBGSKADVNAQDNEDHVPFLH 240  
Db |||||

QY 241 FCSRFEGHDIQVYLLQSDLEVPQHVNIYGDTPHLACYNGKFEVAKEIIQISGTSLETK 300  
Db |||||  
QY 241 FCSRFEGHDIQVYLLQSDLEVPQHVNIYGDTPHLACYNGKFEVAKEIIQISGTSLETK 300  
Db |||||

QY 301 ENIFSETAFHSACTYKSIDLVKFLDQNVININHQGRDGTGLHSACYHGHIRLVQFLL 360  
Db |||||  
QY 301 ENIFSETAFHSACTYKSIDLVKFLDQNVININHQGRDGTGLHSACYHGHIRLVQFLL 360  
Db |||||

QY 361 DNGADMLVACDPSSRSGEKDEQTCMLWAYEKGHDAIVTLKHYKRPQDELPCNEYSQPG 420  
Db |||||  
QY 361 DNGADMLVACDPSSRSGEKDEQTCMLWAYEKGHDAIVTLKHYKRPQDELPCNEYSQPG 420  
Db |||||

QY 421 GDGSYVSPPLGKIKSMTKEKADILLRAGLPSHFHLQISEIFHEIIGSGSFGKVKYG 480  
Db |||||  
QY 421 GDGSYVSPPLGKIKSMTKEKADILLRAGLPSHFHLQISEIFHEIIGSGSFGKVKYG 480  
Db |||||

QY 481 RCRNKIYAIKRYANTYCSKSDVDMFCREVSIICQLNHPCVIOFVGACLDNDPSQFAIVTQ 540  
Db |||||  
QY 481 RCRNKIYAIKRYANTYCSKSDVDMFCREVSIICQLNHPCVIOFVGACLDNDPSQFAIVTQ 540  
Db |||||

QY 541 YISGSLFSLHLEQKRIIDLQSKLIIAVDVAKGMEYLHNLTQPIIHRDLNSHNILLYEDG 600  
Db |||||  
QY 541 YISGSLFSLHLEQKRIIDLQSKLIIAVDVAKGMEYLHNLTQPIIHRDLNSHNILLYEDG 600  
Db |||||

QY 601 HAVVADFGESRFLQSLDEDNMTKQGNLRWMAPEVFTQCTRYTIKADVFSYALCLWEILT 660  
Db |||||  
QY 601 HAVVADFGESRFLQSLDEDNMTKQGNLRWMAPEVFTQCTRYTIKADVFSYALCLWEILT 660  
Db |||||

QY 661 GEIPFAHLKPAADAAADYMHIRPPIGYSTPKPISSLLIRGNACPEGRPEFESEVVMKLE 720  
Db |||||  
QY 661 GEIPFAHLKPAADAAADYMHIRPPIGYSTPKPISSLLIRGNACPEGRPEFESEVVMKLE 720  
Db |||||

QY 721 ECLCNIELMSPASSNSGSLSPSSSDCLVNRGGPGRSHVAALRSRFELEYALNARSYAA 780  
Db |||||  
QY 721 ECLCNIELMSPASSNSGSLSPSSSDCLVNRGGPGRSHVAALRSRFELEYALNARSYAA 780  
Db |||||

QY 781 LSQSAGQYSSQGLSLEEMKRSLOVTPIDKYGVSDPMSSMHFHSNCRSSSFEDESS 835  
Db |||||  
QY 781 LSQSAGQYSSQGLSLEEMKRSLOVTPIDKYGVSDPMSSMHFHSNCRSSSFEDESS 835  
Db |||||

RESULT 2

Q7TOP6

ID Q7TOP6

AC Q7TOP6;

DT 01-OCT-2003 (TRENBLrel. 25, Created)

PRELIMINARY;

PRT; 835 AA.

Db	781	WSQVGHSHNPGLSLEENNRSTQISTVDKYGVSDPMSLTHLHRSQDDSNFDSN	835
RESULT 3			
O17055			
ID	O17055	PRELIMINARY;	PRT; 850 AA.
AC	O17055;		
DT	01-JAN-1998	(TRENBLrel. 05, Created)	
DT	01-OCT-2001	(TRENBLrel. 18, Last sequence update)	
DT	01-OCT-2003	(TRENBLrel. 25, Last annotation update)	
DE	DT	Hypothetical protein.	
DE	C24A1.3		
GN	Caenorhabditis elegans.		
OS	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;		
OC	Rhabditidae; Pelodirinae; Caenorhabditis.		
OX	NCBI_TaxID=6239;		
RA	[1]	SEQUENCE FROM N.A.	
RP	SEQUENCE FROM N.A.		
RP	STRAIN=Bristol N2;		
RP	MEDLINE=99069613; PubMed=9851916;		
RX	None;		
RA	"Genome sequence of the nematode C. elegans: a platform for		
RT	investigating biology. The C. elegans Sequencing Consortium."		
RT	Science 282:2012-2018(1998).		
RL	[2]	SEQUENCE FROM N.A.	
RL	SEQUENCE FROM N.A.		
RC	STRAIN=Bristol N2;		
RC	Connell M.;		
RA	"The sequence of C. elegans cosmid C24A1.1;"		
RT	Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.		
RL	[3]	SEQUENCE FROM N.A.	
RL	SEQUENCE FROM N.A.		
RC	STRAIN=Bristol N2;		
RC	Waterson R.;		
RA	"Direct Submission.;"		
RT	Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.		
RL	EMBL; AF024491; AAB70312.2; -.		
RL	PIR; T32258; T32258.		
DR	HSP; P42773; 1IHB.		
DR	WormPep; C24A1.3; CE27723.		
DR	GO; GO:0005524; F:ATP binding; IEA.		
DR	GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.		
DR	GO; GO:0016740; F:transferase activity; IEA.		
DR	GO; GO:0006469; P:protein amino acid phosphorylation; IEA.		
DR	InterPro; IPR002110; ANK.		
DR	InterPro; IPR000719; Prot_kinase.		
DR	InterPro; IPR001245; Tyr_kinase.		
DR	Fam; PF00023; ank; 9.		
DR	Fam; PF00069; pkinase; 1.		
DR	PRINTS; PR01415; ANKYRIN.		
DR	PRINTS; PR00109; TYRKINASE.		
DR	ProDom; PD000001; Prot_kinase; 1.		
DR	SMART; SM00248; ANK_9.		
DR	PROSITE; PS50088; ANK_REPEAT; 4.		
DR	PROSITE; PS50297; ANK_REPEAT_REGION; 1.		
DR	PROSITE; PS50107; PROTEIN_KINASE_ATP; 1.		
DR	PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.		
KW	Hypothetical protein; ANK repeat; ATP-binding; Kinase; Repeat;		
KW	Transferase; Tyrosine-protein kinase.		
SQ	SEQUENCE 850 AA, 93195 MW; PB844D189B8758A CRC64;		
Query Match	39.1%;	Score 1716.5;	DB 5; Length 850;
Best Local Similarity	41.9%;	Pred. No. 2.5e-127;	
Matches 375;	Conservative 153;	Mismatches 236;	Indels 131; Gaps
QY	1	MGNYSRPTQTCTDEWKKVSESVVITIERLEDDLOIK-----EKELETERL	
Db	1	MGNYSRPLTSCDCLKKKIKSEGVSVRSRLSDVRSRLGWDVQIAAPEKSLDFPK	
QY	48	IFGSDEAFSKVNLNRYTENGSLHLHLCICGG-----KKSHITLML	
Db	60	-----OHMCPENAELKSTOLLSLFHI--ICAGHSDSOPEKLFLOINLPKESI-TLI-	

Qy	91	GURPRLTRNPTALHLAVYKDNALHTSLHSG-ADIQQVGYGGLTALHIAHTIAGHLA	14
Db	110	-----SSQANGFTPLHIAIAYRGDAIYKALIIATKVLDOQSGRHLLPALTALHIAWIGDSEM	165
Qy	150	ADVLLQHGANNVNIQDAVPFTPLHIAAAYGHEQVTRLLLKFCADVNVSGEVDGDRPLHLASA	209
Db	166	L7ILLNSGANIHVDFVHFTALHCATYFGQENAVRTLJISANANLUGAVNDRPIHLAAA	225
Qy	210	KGFLMIAKLMBEGSKADVNAQDNEDHVPFLHFCRFGHHDIVKYLLOSLDEVOPHVV--N	267
Db	226	KGLTITKLLLE--AKADPLLADDEGQALHYAAKSGSLVILANLHKQVGTNRDICARN	283
Qy	268	LYGDTPLHLACVNGKFEVAXEIIQISGTESLTKENIFSETAFHSACVYKGSIDLVRFLLD	327
Db	284	LYGDTALHLSCVSGRLDIVKSIDLSSPTNIVMNVFSEPLHAACTGGKSIELVSFLMK	343
Qy	328	QNVININHGQRDGTGLHSACHYGHIRLVQFLDLNGADMILVA-----CDPSRS	376
Db	344	YFCVDPNYQQDGTALHSACYHGHLEIRVQYLLENGADQSLAGRAFEFGGALRQAQAGFIN	403
Qy	377	SCEK-----DEQTCMLMWAYEKGHDAIVTLTKHY--KRPQ	408
Db	404	RSKVASAIMALNRSDDTSSNASYNSTVSLDDQCTPVIWYERGHDAIVALLKHYAARTV	463
Qy	409	DELPCNEYSPQGGGSGYVSPPLGKTKSMTKEXADILLRAGLPSSHFLQLSEIFHEI	468
Db	464	EGDVCSSEYS--SGESSYTPLESPMGRLTSLTRDKADLLQLRSALPAPFHLCLAEIEFQES	521
Qy	469	IGSGFGKVKYKRCNKIVAKRYRANTYCSKSDVMFCREVSILCOLNHPCVIQFYGAC	528
Db	522	IGSGFGKVKYKTYGKLVAVKRYRANAFGCKSETDMLCREVLSLSLAHPNVVAFVGT	581
Qy	529	LNDPQFAIVTYQISGGSLFLLHEQK---RILDQSLKIADVAKGMGYLH-NLTQPI	584
Db	582	LDDPQFAIITEFVENGSLFRRENGERKNYKYMDFAPFLRAISLDVARGMRYLHESAAPV	641
Qy	585	IHRDLNSHNIILYEDGHAAVADFESRFLOSLDDEDNMTKQGNLRWMAPEVFTQCTRYTI	644
Db	642	IHRDLNSHNIILHAGRSVADDFESRFVQCQEDENLTQKQGNLRWMAPEVFSQSGKYDR	701
Qy	645	KADVFSYALCLWEILTGEIIPFAHLKPAAAAADMAHYHHIRPPI-----GYSIPKPISSLLIR	700
Db	702	KVDVFSFALVIWEIHTABLPFSLKPAAAAEMTYKRGRTPLPNQPTAQFPFALHLSLPQ	761
Qy	701	GNNACPEGRPESEVVMKLEELCNIELMSPASSNSGSLSPSSSDCLVNRGPGCRSHV	760
Db	762	AHPSSSRDPDFVEIVALLE-----PHVESTHTD:SAFST-----V	797
Qy	761	AALRGFE-LEYA-LNARSYAALGSAQGYSSQGLSLEEMKRSLOYTPIKGYIV	813
Db	798	SOLTSQWEQLSVAPPSPASKFPPILSALHGIAATG-TVEELRQR-----IDNNGYV	846
RESULT 4			
Q8BFR0	ID	Q8BFR0	PRELIMINARY; PRT; 165 AA.
AC	Q8BFR0;		
DT	01-MAR-2003	(trEMBLrel. 23, Created)	
DT	01-MAR-2003	(trEMBLrel. 23, Last sequence update)	
DT	01-JUN-2003	(trEMBLrel. 24, Last annotation update)	
DE	Hypothetical ankyrin repeat profile/ankyrin-repeat/ankyrin repeat		
DE	region circular profile/yeast DNA-binding domain containing		
DE	protein.		
GN	D830019J24RIK.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=C57BL/6J; TISSUE=Heart;		
RX	MEDLINE=22354683; PubMed=12466851;		

RA The FANTOM Consortium,  
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 RT 60,770 full-length cDNAs.";  
 RL Nature 420:563-573 (2002).  
 DR EMBL; AK052882; BAC35187.1; -;  
 DR EMBL; AK084817; BAC39285.1; -;  
 DR MGD; MGI:2443276; D830019J24Rik.  
 DR InterPro; IPR002110; ANK.  
 DR Pfam; PF00023; ank; 3.  
 DR SMART; SM00248; ANK; 3.  
 DR PROSITE; PS00088; ANK\_REPEAT; 1.  
 DR PROSITE; PS0297; ANK\_REPEAT\_REGION; 1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 165 AA; 18372 MW; FA9C859545EB657D CRC64;  
 Query Match 15.3%; Score 672; DB 11; Length 165;  
 Best Local Similarity 85.3%; Pred. No. 2.4e-45;  
 Matches 128; Conservative 7; Mismatches 15; Indels 0; Gaps 0;  
 QY 1 MGNYSRPTQCTDEWKKVSESYVITIERLEDDLOIKKELTELNRNIFGSDRAFSKVN 60  
 DB 1 MGNYSRPTQCTSDWKKVSEYVITIERLEDDLOIKKELTELNRNIFGSDRAFSKVN 60  
 QY 61 NYRTENGLSLHLLCCICGGKSHIRTLMLKGLRPSRLTRNGFTALHLAVYKONAE 120  
 DB 61 NYRTERGLSLLHLLCCACGGKSHIRPALMLKGLRPSRLTRNGFPALHLAVYKDSLE 120  
 QY 121 LHSGADIQVGGYGTALHIATIAHLEA 150  
 DB 121 LHSGADVQVGGYGTALHIAAGHPEVS 150  
 RESULT 5  
 ID Q22558 PRELIMINARY; PRT; 546 AA.  
 AC Q22558;  
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Hypothetical protein.  
 GN AT2G17700.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Yamada K., Banh J., Chang C.H., Chang E., Dale J.M., Goldsmith A.D.,  
 RA Lee J.M., Onodera C.S., Quach H.L., Tang C., Toriumi M., Wu H.C.,  
 RA Yamamura Y., Yu G., Yu S., Bowser L., Carninci P., Chen H., Cheuk R.,  
 RA Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,  
 RA Kawai J., Kim C., Lam B., Lin J., Meyers M.C., Miranda M.,  
 RA Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M.,  
 RA Shinn P., Southwick A., Shinozaki K., Davis R.W., Ecker J.R.,  
 RA Theologis A.;  
 RA "Full length cDNA of gene At2g17700 (GI:15227883).";  
 RT Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Yamada K., Banh J., Chan M.M., Chang C.H., Chang E., Dale J.M.,  
 RA Deng J.M., Goldsmith A.D., Lee J.M., Onodera C.S., Quach H.L.,  
 RA Tang C., Toriumi M., Wu H.C., Yamamura Y., Yu G., Bowser L.,  
 RA Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,  
 RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,  
 RA Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,  
 RA Sakurai T., Satou M., Seki M., Shinn P., Southwick A., Shinozaki K.,  
 RA Davis R.W., Ecker J.R., Theologis A.;  
 RA "Arabidopsis Open Reading Frame (ORF) Clones";  
 RT Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
 CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 DR EMBL; AY070086; AAL49781.1; -;  
 DR EMBL; AY096470; AAM20110.1; -;  
 DR PIR; D84555; D84555.  
 DR HSSP; P08631; IAD5.  
 DR GO; GO:0016597; F:amino acid binding; IEA.  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.  
 DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.  
 DR GO; GO:0016740; F:transferase activity; IEA.  
 DR GO; GO:0008152; P:metabolism; IEA.  
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.  
 DR InterPro; IPR002912; ACT.  
 DR InterPro; IPR000719; Prot\_kinase.  
 DR InterPro; IPR008271; Ser\_thr\_kinase.  
 DR InterPro; IPR01245; Tyr\_kinase.  
 DR Pfam; PF01842; ACT; 1.  
 DR Pfam; PF00069; PKinase; 1.  
 DR PRINTS; PR00109; TYRKINASE.  
 DR PRODOM; PD000001; Prot\_kinase; 1.  
 DR PROSITE; PS00011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
 KW Hypothetical protein; ATP-binding; Kinase;  
 KW Serine/threonine-protein kinase; Transferase.  
 SQ SEQUENCE 546 AA; 61509 MW; E74F30F60D91F426 CRC64;  
 Query Match 11.1%; Score 486; DB 10; Length 546;  
 Best Local Similarity 34.4%; Pred. No. 9.9e-30;  
 Matches 114; Conservative 73; Mismatches 120; Indels 24; Gaps 8;  
 QY 416 YSQPGDGSVSPSLGKIKSM--TKEKA-----DILLRAGLPS---HFL 458  
 DB 222 WSQETDGLRDALSKELKLDQPGSKQKSIFFEHDKSSNELIPACIEIPDTGDEWEI 281  
 QY 459 QLSEIEPHEIIGSGSGKVKYKGRNKI VAIKRYRANTYCSKSDVDMFCREVSILCOLNH 518  
 DB 282 DVTQLKTEKKVASSGYDLHRTGYCQEVAKFLKPD-R-VNNEMLREFSQEVFINRKYRH 340  
 QY 519 PCVIQFVAGCLNDPSPFAIVTQVIGSGSLFSLHBEQKRLDLOSKLIITAVDAKMEYLH 578  
 DB 341 KXVVQFLGACTRSPT-LCIVTFEMARGSIYDFLHKQKAFKLTLLKVALDVAKGMSYLH 399  
 QY 579 NLTPILHRLNSHNLVYEDGAVADFGESRFLQSLDEDNNTQPGNLRMAPEVFTQ 638  
 DB 400 Q--NNIHRDLKTANLLMDEHGLVKVADFGVARV--QIESGVMTAETGYRWMAPEVI-E 454  
 QY 639 CTRYTIKADVFSVYALCWEILTEIPFAHLKFAAAADWAYHHIRPPIGYSTPKPISLL 698  
 DB 455 HKPNYHKADVFSYAIWLWELLTGDIPYAFELTPLQAAVGVVQKGLRPKIPKTHPKVKGLL 514  
 QY 699 IRGNWACPEGRPFSEVVMKLEELCNIELM 729  
 DB 515 ERCWHQDPEORPLFEIIEMLQOIMKEVNVV 545  
 RESULT 6  
 ID Q7XXN2 PRELIMINARY; PRT; 603 AA.  
 AC Q7XXN2;  
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Putative serine/threonine-protein kinase ctrl.  
 GN P070581.7.  
 OS Oryza sativa (japonica cultivar-group).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzaceae; Oryza.  
 OX NCBI\_TaxID=39947;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Nipponbare;  
 RA Oliveira A.C., Mattos L.T., Carvalho F.F., Shimano A., Zimmer P.D.,  
 RA Malone G., Dellagostin G.;  
 RT "Oryza sativa nipponbare genomic DNA, chromosome 9, PAC

```
RT clone:P0705B11." ;
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP006548; BAC79157.1; -.
KW Kinase.
SQ
SEQUENCE 603 AA; 66329 MW; 78D406F3E2C3FB12 CRC64;
Query Match 11.0%; Score 482; DB 10; Length 603;
Best Local Similarity 29.4%; Pred. No. 2.4e-29;
Matches 130; Conservative 83; Mismatches 189; Indels 40; Gaps 10;
QY 296 ESLSKNISETAFHSACTYGSIDIVKFLLDQNVININHQGRD-----GHTGLHSA 347
DB 154 EEVASLLNRQSIHPPAFSGSTNLEALALEAS---KSQGDHSDSDNVNVRPMHEI 209
QY 348 CYGHTR-----LVQFLDNGADNMLVACDSRSSEKDEQTCMLWAYEKHDAIVTLL 401
DB 210 TFSITDKPKLLSLLGELGLNIOEAFSTNDGYSLDFVVGWHDDEETDLIESVR 269
QY 402 KHKRQDELPCNEYSQPGDGSYVSPPLGKIKSMTEKADILLRAGLPSHFHQLS 461
DB 270 KE-----IGKIDTQGWSTHSSWSPVENMQIGENSAADHVEIPRDGASWEIDVK 320
QY 462 EIFEHIIGSGFGKYKRCRNKIVAKEYRANTYCSKSDVM---PCREVSILCOLNH 518
DB 321 LLKFGNKVAGSGYGLRYGTYCSQDVAIKVLP---ERINADMQREFAQEYVIMRKVRH 376
QY 519 PCVIOFVGACLNDPSQFAIVTQVIGSGSLFSLHLEQKRLDLOSKLLIADVAVAKGMEYLH 578
DB 377 KNVVQFIGACTKEPN-LCIVETVMGSGSYDYLHKHGVFKLPALLGVVMDVSKGMSYLH 435
QY 579 NLTPQIIRHLDNHNILLYEDGHAVVADFGESRFLQSLDDEDMTKQGNLRMAPEFTQ 638
DB 436 Q--NNIHRDLKTNLLMDENGTVKVADFGARV--KAQSGVMTAETGYRMAPEVI-E 490
QY 639 CTRYTKADVSLVALCWEILGELIPEHLKPAARAAADWAYHRIIPPGYSTPKPISSLL 698
DB 491 HKPYDHKADVFSEGLMWELLTKIPYEYLTQRAVGVVQKGLRPTTPKNAHAKLSLL 550
QY 699 IRGNWACPEGRPEFSEVVMKLE 720
DB 551 QKQVQEPAPRPFDSILETLQ 572
RESULT 7
Q9LMF8 PRELIMINARY; PRT; 445 AA.
ID Q9LMF8
AC Q9LMF8
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE F16A14.22.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN 1.
SEQUENCE FROM N.A.
RA Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C., Khan S.,
RA Kim C., Altafi H., Bei Q., Chin C., Chiou J., Choi E., Conn L.,
RA Conway A., Gonzales A., Hansen N., Howing B., Koo T., Lam B., Lee J.,
RA Lenz C., Li J., Liu A., Liu K., Liu J., Liu S., Mukharsky N., Nguyen M.,
RA Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaveri A.,
RA Toriumi M., Vaysberg M., Yu G., Federspiel N., Eckert J.R.;
RA "Genomic sequence for Arabidopsis thaliana BAC F16A14 from chromosome
RT I.",
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
RN 2.
SEQUENCE FROM N.A.
RA Eckert J.R.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
RN 3.
```

```
RP SEQUENCE FROM N.A.
RA Eckert J.R.;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
RN 4.
SEQUENCE FROM N.A.
RA Cheuk R., Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C.,
RA Khan S., Kim C., Altafi H., Bei Q., Chin C., Chiou J., Choi E.,
RA Conn L., Conway A., Gonzales A., Hansen N., Howing B., Koo T., Lam B.,
RA Lee J., Lenz C., Li J., Liu A., Liu K., Liu J., Liu S., Mukharsky N.,
RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,
RA Thaveri A., Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N.,
RA Theologis A., Eckert J.R.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AC068197; RAF79405.1; -.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00023; ank; 1.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00248; ANK; 1.
DR SMART; SM00219; TYRK; 1.
DR PROSITE; PS50088; ANK_REPEAT; 1.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR ANK repeat; ATP-binding; Kinase; Repeat;
KW Serine/threonine-protein kinase; transferase.
SQ
SEQUENCE 445 AA; 50171 MW; EAAF39E47C1AD2E6 CRC64;
Query Match 10.9%; Score 476.5; DB 10; Length 445;
Best Local Similarity 32.6%; Pred. No. 4.1e-29;
Matches 134; Conservative 79; Mismatches 153; Indels 45; Gaps 13;
QY 322 VKFLDQNVININHQGRDGHGLHSAHYGHRLVQFLDNGADNMLVACDPSRSSEK 381
DB 53 VRKLEEDPTLVHARDYDKRTTLHVASLHGWIDVVKCLLEFGADVN-----AQDRW 103
QY 382 EQTCMLWAYEKGHDAIVTLLKHKRPQDELPCNYSQPGDGSYVS---VPSPLGKIKSM 438
DB 104 KNTPLADAEGARKQKXMIELKSHGGLS---YLSKFTMQQNGSHPEPKVPPIPK--- 156
QY 439 TKSKADILLRAGLPSHFHQLSEIEFHE--IGSGSGKVKYKRCRNKIVAKEYRANT 496
DB 157 ---KCD-----WEIEPAELDFSNAAITGKSGFGEIVKAYWRGTTPAVKRILFSL 202
QY 497 YCSKSDVDMFCREVSTLCOINHPCVIQVYGACINDPSQFAIVTQVIGSGSLFSLHLEQKR 556
DB 203 SDRRLVIQDFRHEVDLLVLRHFNIVQFLGA-VTERKPLMLITEYLRGGDLHQLYKE-KG 260
QY 557 ILDLQSKLLIADVAVAKGMEYLHNLTQPIIHRDNLGNHLLYEDG--HAVVADFGSRFLQ 614
DB 261 GLTPTTAVNPALDIASGMTYLHNEPNVILHRDLKPRNVLNVSADHLKVGDFGLSKLIK 320
QY 615 ---SLDEDNMTKQGNLRMAPEVFTQCTRYTIKADVSVALCWEILTGETIPFAHLKPA 671
DB 321 VQNSHDVYKMTGTGTSYRYNAPEVFKH-RRYDKKVDVFSFAMILEMELEGEPFANHEPY 379
QY 672 AAAADWAYHRIIPPGYSTPKP-ISSLIRGNWACPEGRPEFSEVVMKLEE 721
DB 380 EAAKHVSDGH-RPTFRSKGCTPDRLRLIVKCDADNMNQPSFDLILKLEK 429
RESULT 8
```



```
Db 334 VDFGLSKIIKAHANDVYKMTGETGSRYMAPEVFKH-RKYDKKVDVFSFAMILYEMLE 392
QY 661 GEIPFAHLPAAPAAADMAVHHIRPPT--GYSGPKPISSLIRGNACPEGRPFSEVVM 717
Db 393 GDFPFNFPYERAAKYVGGSH-RPFERSKFT--NELKELIELCWSGDIHLRPSLEILK 449
QY 718 KLEE 721
Db 450 RLEK 453

RESULT 10
Q9X187 QX187 PRELIMINARY; PRT; 438 AA.
AC Q9X187
DT 01-NOV-1999 (T-EMBLrel. 12, Created)
DT 01-NOV-1999 (T-EMBLrel. 12, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE F7A19.9 protein.
GN F7A19.9.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP STRAIN=cv. Columbia;
RC Federspiel N.A., Palm C.J., Conway A.B., Conn L., Hansen N.F.,
RA Altafi H., Araujo R., Huizar D., Rowley D., Buehler E., Dunn P.,
RA Gonzalez A., Kremetskaia I., Kim C., Lenz C., Li J., Liu S.,
RA Luros S., Schwartz J., Shinn P., Toriumi M., Vysotskaia V.S.,
RA Walker M., Yu G., Ecker J., Theologis A., Davis R.W.;
RA Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
RL EMBL; AC007576; AAD39286.1; -.
DR PIR; C86273; C86273.
DR HSP; Q00534; 1318.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR008271; Ser_Thr_pkin_AS.
DR Pfam; PF00023; ank; 1.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00248; ANK; 1.
DR SMART; SM00219; TyrKc; 1.
DR PROSITE; PS50088; ANK_REPEAT; 1.
DR PROSITE; PS50297; ANK_REPEAT; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR ANK repeat; ATP-binding; Kinase; Repeat;
KW Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 438 AA; 49335 MW; C61F80724E24BBD2 CRC64;
```

```
Query Match 10.8%; Score 475; DB 10; Length 438;
Best Local Similarity 32.8%; Pred. No. 5.2e-29;
Matches 134; Conservative 73; Mismatches 155; Indels 46; Gaps 12;

QY 322 VKFLDQNVINHGQRDGTGLHSACVHGIRLVQFLDNGADMVLVACDPSRSSGEKD 381
Db 53 VRKLEEDPTLVHARDYDKRTPHVASLHGWDVVKCLLEFGADV-----AQDRW 103
QY 382 EOTCLMAYEKGHDAIVTLKHKKRPQDELPCNYSQPGDGSVSPSLGKIKMTKE 441
```

```
Db 104 KNTPLADAGARKQKWIILLKSHGG-----LSYGQNGSHFEKPVPPPIPK----- 149
QY 442 KADILLRAGLSHFLQLSEIEPHE--IIGSGFGKVKYKCRNKIVAKRYEANTYCS 499
Db 150 KCD-----WEIEFAELDFSAAMIKGSGFEIVKAYWRGTPVAVKRIILPSLSD 198
QY 500 KSDVDMFCREVSILCOLNHPCVIQFVGACLNDDPSQFAIVTQVISGSLFSLHFKRILD 559
Db 199 RLVIQDFRHEVDLLVLRHPIVQFLGA-VTERKPLMLITEYLRGGDLHQYLKE-KGLT 256
QY 560 LQSKLIADVAKGMEYLNLTQPIIHRDLNSHNTLLYEDG--HAVVADFQESRPLQ--- 614
Db 257 PTTAVNFALDIARGMTYLNHPNVIIHRDLKPRNVLVNSSADHLKVGDFGLSKLIQVN 316
QY 615 SLDEDMTKQPGNLRMAPEVFTQCTRYTIKADVSVALCWEILTGEIPFAHLKPAAAA 674
Db 317 SHDVYKMTGETGSRYMAPEVFKH-RRYDKKVDVFSFAMILYEMLEGEPPFANHEPYEA 375
QY 675 ADMAYHHIRPPIGYSIPKP-ISSLIRGNACPEGRPFSEVVMKLEE 721
Db 376 KHVSDGH-RPTFRSKGCTPDLRELIVKWDADMNQRPSPFLDLKLEK 422

RESULT 11
Q8S9J9 Q8S9J9 PRELIMINARY; PRT; 438 AA.
AC Q8S9J9
DT 01-JUN-2002 (T-EMBLrel. 21, Created)
DT 01-JUN-2002 (T-EMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE A1G14000/F7A19.9.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Shinn P., Chen H., Cheuk R., Kim C.J., Meyers M.C., Banh J.,
RA Bowser L., Carninci P., Chang E., Dale J.M., Goldsmith A.D.,
RA Hayashizaki Y., Iehida J., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin J., Miranda M., Narusaka M.,
RA Nguyen M., Onodera C.S., Palm C.J., Quach H.L., Sakurai T., Satou M.,
RA Seki M., Southwick A., Tang C.C., Toriumi M., Wu H.C., Yamada K.,
RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
RA Ecker J.R.;
RA "Arabidopsis cDNA clones."
RT Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
RL -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC EMBL; AY075653; AAL77660.1; -.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR008271; Ser_Thr_pkin_AS.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00023; ank; 1.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00248; ANK; 1.
DR SMART; SM00219; TyrKc; 1.
DR PROSITE; PS50088; ANK_REPEAT; 1.
DR PROSITE; PS50297; ANK_REPEAT; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR ANK repeat; ATP-binding; Kinase; Repeat;
KW Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 438 AA; 49363 MW; 396FDA64886064A CRC64;
```



RA "Arabidopsis cDNA clones";  
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.  
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
DR EMBL; AF462860; AAL58946.1; -.  
DR GO; GO:000524; F:ATP binding; IEA.  
DR GO; GO:0004674; P:protein serine/threonine kinase activity; IEA.  
DR GO; GO:0004713; P:protein-tyrosine kinase activity; IEA.  
DR GO; GO:0016740; F:transferase activity; IEA.  
DR GO; GO:0004688; P:protein amino acid phosphorylation; IEA.  
DR InterPro; IPR000719; Prot\_kinase.  
DR Ecker J., Theologis A., Davis R.W.;  
DR InterPro; IPR008271; Ser\_thr\_pkin\_AS.  
DR Pfam; PF00069; pkinase; 1.  
DR PRINTS; PD00001; Prot\_kinase; 1.  
DR PROSITE; PS00011; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.  
SQ SEQUENCE 525 AA; 58841 MW; 88C2A6D6C6913856 CRC64;  
  
Query Match 10.8%; Score 472; DB 10; Length 525;  
Best Local Similarity 32.4%; Pred. No. 1.2e-28;  
Matches 122; Conservative 68; Mismatches 127; Indels 60; Gaps 13;  
  
Qy 430 SPGLKTKSM-----TKEKADI-----LLLRAG-----LPSHFHLQSLSEIEFHEIIG 470  
Db 155 SPSPGPIRLGLTRIQERVKSKDTGWSKLPDNTGRRVSAVEASEFRVDMSKLFFGLKFA 214  
  
Qy 471 SGSGFKYKRCNKIVAKRYA-----NTYKSDVD-MFCREVSILCOLNHPVCVIOFV 525  
Db 215 HGLYSRLHYKEDKAVAKLITVPDDONGCUGARLEKQFTKEVTLRLTHPNVIRFV 274  
  
Qy 526 GACINDPSQFAIVTQVIGSGSLSLHE-QKRILDLQSKLIIADVAKGMEYDHLNLTQPI 584  
Db 275 GA-YKDPVPVCLTVQLPEGLSLASFLHKPENRSLPLKLIIEFVIDIARGMEYIH--SRRI 331  
  
Qy 585 IHRDLNSHNLIIYDGHVADFG-----ERFLQSDENMTQPGNLRMAPEVTQCTR 641  
Db 332 IHRDLKPENLIDEEFHLKIADFGIACEEYICDMLADD-----PGTYRMAPEMIKR-KP 385  
  
Qy 642 YTIKADYFVALCLWEITLGEIPFAHLKPAADAAADWAYHHIRPPTGYSTPKPISSILLIRG 701  
Db 386 HGRKADYFGLVNLWNAAGAIYEDYNFIOAFAVVKHNIKPAIFGDCPVAMKALIEQC 445  
  
Qy 702 WNACPEGRPFSEVVMKLECLNIELMSPASSNSGSLSPSSSDCL----- 749  
Db 446 WSVAPDKRPFQWIKVLEQFAISLE-----RKGNLNLSSKICKDPRKGLKHQWK 497  
  
Qy 750 ---VNRGPGGRSHVAAAL 763  
Db 498 LQPVHAGGGGGSSSGL 514  
  
RESULT 14  
Q8RWL6 PRELIMINARY; PRT; 570 AA.  
ID Q8RWL6  
AC Q8RWL6;  
DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Hypothetical protein.  
GN AT4G35780  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC eucosids II; Brassicales; Brassicaceae; Arabidopsi.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Southwick A., Karlin-Neumann G., Nguyen M., Lam B., Miranda M.,  
RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,  
RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,  
RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,  
RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,  
RA Ecker J., Theologis A., Davis R.W.;  
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Tripp M., Southwick A., Karlin-Neumann G., Nguyen M., Miranda M.,  
RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,  
RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,  
RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,  
RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,  
RA Ecker J., Theologis A., Davis R.W.;  
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.  
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
DR EMBL; AY03017; AAM13016.1; -.  
DR EMBL; AY128938; AAM91338.1; -.  
DR GO; GO:0016597; F:amino acid binding; IEA.  
DR GO; GO:000524; F:ATP binding; IEA.  
DR GO; GO:0004674; P:protein serine/threonine kinase activity; IEA.  
DR GO; GO:0004713; P:protein-tyrosine kinase activity; IEA.  
DR GO; GO:0016740; F:transferase activity; IEA.  
DR GO; GO:0008152; P:metabolism; IEA.  
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.  
DR InterPro; IPR002912; ACT.  
DR InterPro; IPR000719; Prot\_kinase.  
DR InterPro; IPR008271; Ser\_thr\_pkin\_AS.  
DR InterPro; IPR001245; Tyr\_pkinase.  
DR Pfam; PF01842; ACT; 1.  
DR Pfam; PF00069; pkinase; 1.  
DR PRINTS; PR00109; TYRKINASE.  
DR PRODOM; PD000001; Prot\_kinase; 1.  
DR PROSITE; PS00011; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
KW Hypothetical protein; ATP-binding; Kinase;  
KW Serine/threonine-protein kinase; Transferase.  
SQ SEQUENCE 570 AA; 64696 MW; D0961D218BE49344 CRC64;  
  
Query Match 10.7%; Score 469.5; DB 10; Length 570;  
Best Local Similarity 32.2%; Pred. No. 2.2e-28;  
Matches 112; Conservative 74; Mismatches 120; Indels 42; Gaps 10;  
  
Qy 388 WAYEKGHDAIVTLKKHYKRPQDELPCNEYSQPGDGSVYVSPSLGKIKSMT-----REK 442  
Db 228 WSQETETGLKDALKKIRFKDQ-PCS-----KQKSIFFPEHDKST 267  
  
Qy 443 ADILLRAGLPS-----HFHLQSLSEIEFHEIIGSGFGKYKRCNKIVAKRYANYC 498  
Db 268 NELLPAACEIPTDGTDEWIDMKQLKIEKKVACGSYGELEFRGTCSEVAILKLP- 323  
  
Qy 499 SKSDVDM---FCREVSILCOLNHPCVIOFVQACLNDDPSQFAIVTQVIGSGSLFSLHEOK 555  
Db 324 ERVNAEMLRFSQEVYIMRKVHKVNVQFICACTSPN-LCIVTEFTGSIYDFLKHK 382  
  
Qy 556 RILDQSKLIIADVAKGMEYDHLNLTQPIIHRDLNSHNLIIYDGHVADFGSRFLOS 615  
Db 383 GVFKIQSLKVALDVSQGMNLLHQ--NNIHRDLKTANLLMDEHEVVKVADFGVARV--Q 438  
  
Qy 616 LDEDNMTQPGNLRMAPEVTQCTEYTIKADYFVALCLWEITLGEIPFAHLKPAADAA 675  
Db 439 TEGSVMTAETGTYYRWNAPEVI-EHKYDHRADVFSYVAILVWELLGELPYSYLTPLQAV 497  
  
Qy 676 DMAYHHRPPIGYSIPKPISSLLIRGNACPEGRPFSEVVMKLECL 723  
Db 498 GVQVKGLRPKIPKETHPKLTLELKEKQDQDPAIRPNFAEIIEMLNQLI 545  
  
RESULT 15  
Q86AT8 PRELIMINARY; PRT; 638 AA.  
ID Q86AT8  
AC Q86AT8;  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Similar to Dictyostelium discoideum (Slime mold). ankyrin repeat

RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,  
RA Ecker J., Theologis A., Davis R.W.;  
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Tripp M., Southwick A., Karlin-Neumann G., Nguyen M., Miranda M.,  
RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,  
RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,  
RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,  
RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,  
RA Ecker J., Theologis A., Davis R.W.;  
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.  
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
DR EMBL; AY03017; AAM13016.1; -.  
DR EMBL; AY128938; AAM91338.1; -.  
DR GO; GO:0016597; F:amino acid binding; IEA.  
DR GO; GO:000524; F:ATP binding; IEA.  
DR GO; GO:0004674; P:protein serine/threonine kinase activity; IEA.  
DR GO; GO:0004713; P:protein-tyrosine kinase activity; IEA.  
DR GO; GO:0016740; F:transferase activity; IEA.  
DR GO; GO:0008152; P:metabolism; IEA.  
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.  
DR InterPro; IPR002912; ACT.  
DR InterPro; IPR000719; Prot\_kinase.  
DR InterPro; IPR008271; Ser\_thr\_pkin\_AS.  
DR InterPro; IPR001245; Tyr\_pkinase.  
DR Pfam; PF01842; ACT; 1.  
DR Pfam; PF00069; pkinase; 1.  
DR PRINTS; PR00109; TYRKINASE.  
DR PRODOM; PD000001; Prot\_kinase; 1.  
DR PROSITE; PS00011; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
KW Hypothetical protein; ATP-binding; Kinase;  
KW Serine/threonine-protein kinase; Transferase.  
SQ SEQUENCE 570 AA; 64696 MW; D0961D218BE49344 CRC64;  
  
Query Match 10.7%; Score 469.5; DB 10; Length 570;  
Best Local Similarity 32.2%; Pred. No. 2.2e-28;  
Matches 112; Conservative 74; Mismatches 120; Indels 42; Gaps 10;  
  
Qy 388 WAYEKGHDAIVTLKKHYKRPQDELPCNEYSQPGDGSVYVSPSLGKIKSMT-----REK 442  
Db 228 WSQETETGLKDALKKIRFKDQ-PCS-----KQKSIFFPEHDKST 267  
  
Qy 443 ADILLRAGLPS-----HFHLQSLSEIEFHEIIGSGFGKYKRCNKIVAKRYANYC 498  
Db 268 NELLPAACEIPTDGTDEWIDMKQLKIEKKVACGSYGELEFRGTCSEVAILKLP- 323  
  
Qy 499 SKSDVDM---FCREVSILCOLNHPCVIOFVQACLNDDPSQFAIVTQVIGSGSLFSLHEOK 555  
Db 324 ERVNAEMLRFSQEVYIMRKVHKVNVQFICACTSPN-LCIVTEFTGSIYDFLKHK 382  
  
Qy 556 RILDQSKLIIADVAKGMEYDHLNLTQPIIHRDLNSHNLIIYDGHVADFGSRFLOS 615  
Db 383 GVFKIQSLKVALDVSQGMNLLHQ--NNIHRDLKTANLLMDEHEVVKVADFGVARV--Q 438  
  
Qy 616 LDEDNMTQPGNLRMAPEVTQCTEYTIKADYFVALCLWEITLGEIPFAHLKPAADAA 675  
Db 439 TEGSVMTAETGTYYRWNAPEVI-EHKYDHRADVFSYVAILVWELLGELPYSYLTPLQAV 497  
  
Qy 676 DMAYHHRPPIGYSIPKPISSLLIRGNACPEGRPFSEVVMKLECL 723  
Db 498 GVQVKGLRPKIPKETHPKLTLELKEKQDQDPAIRPNFAEIIEMLNQLI 545  
  
RESULT 15  
Q86AT8 PRELIMINARY; PRT; 638 AA.  
ID Q86AT8  
AC Q86AT8;  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Similar to Dictyostelium discoideum (Slime mold). ankyrin repeat

DE containing protein.  
 OS Dictyostelium discoideum (Slime mold).  
 CC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.  
 OX NCBI\_TaxID=44889;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=AX4;  
 RX MEDLINE=22092622; PubMed=12097910;  
 RA Gloeckner G., Eichinger L., Szafranski K., Pachebat J., Dear P.,  
 RA Lehmann R., Baumgart C., Parra G., April J.F., Guigo R., Kumpf K.,  
 RA Tungal B., Cox B., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.;  
 RT "Sequence and analysis of chromosome 2 of Dictyostelium discoideum.";  
 RL Nature 418:79-85(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=AX4;  
 RC STRAIN=AX4;  
 RA Baumgart C.;  
 RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AC116330; AA051046.1; -;  
 DR GO; GO:0005524; P:ATP binding; IEA.  
 DR GO; GO:0004674; P:protein serine/threonine kinase activity; IEA.  
 DR GO; GO:0004713; P:protein-tyrosine kinase activity; IEA.  
 DR GO; GO:0005468; P:protein amino acid phosphorylation; IEA.  
 DR InterPro; IPR002110; ANK.  
 DR InterPro; IPR000719; Prot\_Kinase.  
 DR InterPro; IPR001660; SAM.  
 DR InterPro; IPR002290; Ser\_thr\_kinase.  
 DR InterPro; IPR008371; Ser\_thr\_pkin\_AS.  
 DR InterPro; IPR001245; Tyr\_kinase.  
 DR Pfam; PF00023; ank; 5.  
 DR Pfam; PF00069; pkinase; 1.  
 DR Pfam; PF00536; SAM; 1.  
 DR PRINTS; PR00109; TYRKINASE.  
 DR ProDom; PD000001; Prot\_kinase; 1.  
 DR SMART; SM00248; ANK; 6.  
 DR SMART; SM00454; SAM; 1.  
 DR SMART; SM00220; S\_TKc; 1.  
 DR SMART; SM00219; Tyrc; 1.  
 DR PROSITE; PS00088; ANK\_REPEAT; 3.  
 DR PROSITE; PS0297; ANK\_REPEAT\_REGION; 1.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE; PS00111; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
 DR PROSITE; PS00105; SAM\_DOMAIN; 1.  
 DR PROSITE; PS00105; SAM\_DOMAIN; 1.  
 SQ SEQUENCE 638 AA; 72983 MW; 4897A1A6EA62194A CRC64;

Query Match 10.7%; Score 469.5; DB 5; Length 638;  
 Best Local Similarity 25.5%; Pred. No. 2.6e-28;  
 Matches 178; Conservative 111; Mismatches 266; Indels 143; Gaps 23;

QY 90 KGLRPSRLT---RNGFTALHLYAVKDAELITSLH-----SGADIQQVGGTALHAT 142  
 DB 30 KKISKSNLTSPDQYQCSALTALKNNNEEMVELLSLCVTLKADINTFDKNGPSALHQA 89  
 QY 143 TAGHLEAADVLILQHGANVNIQDAVEFTPLHIAAY-----GHEQVTRLILKFGADV 195  
 DB 90 SSDDLRLKRVLYQENINVDVQNDLNTPIH---YFCQKFRSPNCQBPFLFIQKGVNVA 146  
 QY 196 SGEVGRDLPHLASAGFNIAKLMEG-----SKADVNAQDNEDHVPFLFCSPFGHDI 251  
 DB 147 QNKGTEFLH---KAIFNNSVRLMMVGLLKNGANVNLATQFQSPGLHYAVELGREDLV 202  
 QY 252 KYLLQSLDLEQVPHVNNYGDTPHLACVNGKFEVAKIIQISGTESLTKENIFSETAFHS 311  
 DB 203 SVLLKAGADV-CVGTKEKTPYQLAVEEGNKMTARIKKY-----KDLFDWLQKHG 253  
 QY 312 ACTYCKSLDKVFLIDQ-----NVINIHQGRDCHTGLHSACHGHIRLVQFLJLN 362  
 DB 254 FEQYKDAFLKEEMFLDELGEMSEDLNKMGITSTGT-----RLRIK----- 295  
 QY 363 GADNMLVACDPSRSSGKDEQTCMLWAVEKGHDAIVLLKHYKRPQ-DELPCEYSPGG 421  
 DB 296 -----ETSNLANEQT-----KKPKTPELIIE----- 317

QY 422 DGSYVSPSLGKIKSMTEKADILLRAGLPSPHH-----LQISEIEFFHIIIGSGSPG 475  
 DB 318 -----DPTP-----PTPDISGLRHSLSHLRHVGEVNIINDNELEYTEKLGAGSSG 363  
 QY 476 KYKGRCRNKIVAKRYRANTYCSKSDVMFCREVSILCOLNHPCVIQVFGACLNDPSQF 535  
 DB 364 KYKGLYRGKEVAIKVLKSMT--ESKEIEBEFKKEFQIMSAIRSKHVHVEFGAVLE--PKL 419  
 QY 536 AIVTQYISGGLPSLLHEQKRIILDQSKLIIADVAKGMEYLHNLTPQIIHRDLASHNLL 595  
 DB 420 CMVMENCSRGLYHVMNNDLDIGWERTFRFAIETVRGIECLKWDPPIVHRDLKSLNLL 479  
 QY 596 LYEDGHAVADVDFGESRFLOSLDEDNMTKQGNLRWMAPEVFTQCTRYTIKADVFSYALCL 655  
 DB 480 VNDKWEIKVCDFFGLSRFTGNSLETLVKMRGTFCAYCAPEVY-YGEQFSGKSDVYSIAVIL 538  
 QY 656 WEILTG-----EIPFAHLK-----PAAADADNAYHHIRPPIGYSIPKPISSILLIRGNAC 705  
 DB 539 WELVTRCINGRYERPFSEYKNLQHFQIIQTAKNLRPTIPNACPESLVSLIQCDWDPN 598  
 QY 706 PEGRPEFSEVVMKL---BECLCNIE---LMSPASSN 735  
 DB 599 LENRPTCTDILSRVLTIEYRSNIQTWNLLIVLPFN 636

Search completed: September 2, 2004, 14:24:13  
 Job time : 153 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 2, 2004, 14:08:16 ; Search time 27 Seconds  
(without alignments)  
1610.318 Million cell updates/sec

Title: US-10-626-173-2  
Perfect score: 4390  
Sequence: 1 MGNYKSRPTQTCTDEWKVV.....PMSSMHFHSRNSSPEDSS 835

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	465	10.6	1584	1 KYK1_DICDI	P18160 dictyosteli
2	460.5	10.5	1880	1 ANK1_HUMAN	P16157 homo sapien
3	457.5	10.4	1862	1 ANK1_MOUSE	Q02357 mus musculu
4	450.5	10.3	410	1 KYK1_DICDI	P18161 dictyosteli
5	441	10.0	1059	1 Y379_HUMAN	O15084 homo sapien
6	435	9.9	3924	1 ANK2_HUMAN	Q01484 homo sapien
7	427.5	9.7	4377	1 ANK3_HUMAN	Q12955 homo sapien
8	419	9.5	579	1 M3K7_MOUSE	Q62073 mus musculu
9	419	9.5	606	1 M3K7_HUMAN	Q43318 homo sapien
10	414	9.4	954	1 M3KA_HUMAN	Q02779 homo sapien
11	404.5	9.2	821	1 CTR1_ARATH	Q05609 arabidopsis
12	401	9.1	757	1 HT16_HYDAT	P53356 hydra atten
13	400.5	9.1	452	1 ILK1_HUMAN	Q13418 homo sapien
14	396	9.0	452	1 ILK_MOUSE	O55222 mus musculu
15	390.5	8.9	452	1 ILK3_HUMAN	P57043 homo sapien
16	388	8.8	451	1 ILK_CAVPO	P57044 cavia porce
17	375	8.5	832	1 ANR3_HUMAN	P57078 homo sapien
18	367.5	8.4	625	1 ITK_MOUSE	Q03526 mus musculu
19	364	8.3	620	1 ITK_HUMAN	Q08881 homo sapien
20	358.5	8.2	467	1 MATK_RAT	P41243 rattus norv
21	357.5	8.1	822	1 FER_HUMAN	P16591 homo sapien
22	356	8.1	859	1 M3KC_HUMAN	Q12852 homo sapien
23	356	8.1	888	1 M3KC_MOUSE	Q60700 mus musculu
24	355	8.1	394	1 M3K9_HUMAN	P80192 homo sapien
25	354	8.1	768	1 YB23_HUMAN	Q9ulj7 homo sapien
26	354	8.1	888	1 M3KC_RAT	Q63796 rattus norv
27	353	8.0	509	1 STK_HYDAT	P17713 hydra atten
28	353	8.0	735	1 RNSA_MOUSE	Q05921 mus musculu
29	352.5	8.0	1520	1 ABL_DROME	P00522 drosophila
30	351	8.0	647	1 RAF1_CHICK	P05625 gallus gall
31	349.5	8.0	505	1 MATK_MOUSE	P41242 mus musculu
32	349	7.9	450	1 CSK_CHICK	P41239 gallus gall
33	348.5	7.9	638	1 RAF1_XENLA	P09560 xenopus lae

34	346.5	7.9	552	1 SR64_DROME	P00528 drosophila
35	345.5	7.9	648	1 RAF1_RAT	P11345 rattus norv
36	345	7.9	505	1 SRK1_SPOLA	P42686 sporgilla 1
37	344	7.8	380	1 MIL_AVIMH	P00531 avian retro
38	342.5	7.8	648	1 RAF1_MOUSE	Q99n57 mus musculu
39	342	7.8	539	1 RIK2_MOUSE	P58801 mus musculu
40	341	7.8	1442	1 DAK1_MOUSE	Q80ye7 mus musculu
41	340	7.7	1122	1 TIE2_MOUSE	Q02858 mus musculu
42	340	7.7	1124	1 TIE2_HUMAN	Q02763 homo sapien
43	340	7.7	1125	1 TIE2_BOVIN	Q06807 bos tauris
44	339	7.7	1432	1 DAK1_HUMAN	P53355 homo sapien
45	337	7.7	527	1 TXK_HUMAN	P42681 homo sapien

ALIGNMENTS

RESULT 1

KYK1_DICDI	ID	KYK1_DICDI	STANDARD;	PRT;	1584 AA.
AC	P18160;				
DT	01-NOV-1990 (Rel. 16, Created)				
DT	01-OCT-1996 (Rel. 34, Last sequence update)				
DT	15-MAR-2004 (Rel. 43, Last annotation update)				
DE	Non-receptor tyrosine kinase spore lysis A (EC 2.7.1.112) (Tyrosine-protein kinase 1).				
DE	SPLA OR PYKA OR DPVK1.				
GN	Dictyostelium discoideum (Slime mold).				
OS	Dictyostelium discoideum (Slime mold).				
OC	Eukaryota, Mycetozoa; Dictyostelida; Dictyostelium.				
OX	NCBI_TaxID=44689;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=JH10;				
EX	MEDLINE=97053827; PubMed=8898241;				
EA	Nuckolls G.H., Osherv N., Loomis W.F., Spudich J.A.;				
RT	"the Dictyostelium dual-specificity kinase spla is essential for spore differentiation.";				
RL	Development 122:3295-3305(1996).				
RN	[2]				
RP	SEQUENCE OF 1248-1584 FROM N.A.				
EX	MEDLINE=90287147; PubMed=1972546;				
RA	Tan J.L., Spudich J.A.;				
RT	"Developmentally regulated protein-tyrosine kinase genes in Dictyostelium discoideum.";				
RL	Mol. Cell. Biol. 10:3578-3583(1990).				
CC	!- FUNCTION: Essential for spore differentiation.				
CC	!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein tyrosine phosphate.				
CC	!- DEVELOPMENTAL STAGE: Expressed throughout development with a peak during the mound stage of morphogenesis.				
CC	!- SIMILARITY: Belongs to the Tyr family of protein kinases.				
CC	!- SIMILARITY: Contains 1 sterile alpha motif (SAM) domain.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).				
CC	-----				
EMBL	U32174; BAB41125.1; ..				
EMBL	M33785; AAA33202.1; ..				
PIR	T18276; T18276.				
DictyBase	DDB0001716; splA.				
InterPro	IPR000719; Prot_kinase.				
InterPro	IPR001660; SAM.				
InterPro	IPR003877; SPRY_receptor.				
InterPro	IPR001245; Tyr_Pkinase.				
InterPro	IPR008266; Tyr_Pkinase_AS.				
Pfam	PF00069; pkinase; 1.				
Pfam	PF00536; SAM; 1.				
Pfam	PF00622; SPRY; 3.				



251 VIMVRLLDRCAGAQIETKTDLPFLCAARN---GHVRISSEILLDHGAPIAKTKNGLSLP 307  
 71 LHLG-----CI-----CGKKSHIRTLMLKGURPSRL 97  
 308 IHMAAQSDHLDCVLLQLQYDAEIDITDLHPLFLHVAACHGHRV-AKVLLDKGAKPNR 366  
 98 TENGFTALHVAAYKDNAAELITSLHSGADIQQGVGGLTALHIAITIAHLEAADVLLQHG 157  
 367 ALNGFTPLHIAACKXHVVRWELLKTKGASIDAVTESGLTPLHVASFMGHLPIVKNLLQRG 426  
 158 ANVNIQDAVFTPLHIAAYYGHQVTRLKLLKFGADVNSGEVGDRLPLHASAKGFLNTAK 217  
 427 ASPNSVNVKVTPLHMAARAGHTEVAKYLLQNKAKYNARAKDDQTPFLHCAARIGHNTNMVK 486  
 218 LIMBEGSKADVNAQDNEDHVPFLHFCRFGH-----HDIVK 252  
 487 LLLLE--NNANFNLTATGHTPLHIAAREGHVETVLALLEKASQACMTKKGFTPLHVAAK 544  
 253 Y-----LLQSD-----LEV-----QPHVVNIYGDTPT 273  
 545 YGKVRVAELLERDAHFNAAAGKNGLTPLHVAVHHNNLDIVKLLPRGGSPHSPAMNGYTP 604  
 274 LHLACYNKCPVAKEIITQISGTSLETKENIFSTAFHSACTYKSIDLVKFLDQNV--- 330  
 605 LHIAACKQNVARSLLQYGG--SANAESQGVTPPLHAAQEGHA-EWVALLLSKQANGN 661  
 331 -----ININHQGRDGHGTLHLSACYGHGIRLVQVFLLD 361  
 662 LGNKSGLTPLHLVAQEGHVPVADVLIKHGVMVDATTMRGYTPLHVAHYGNIKLVKFLQ 721  
 362 NGADNNLVACDPSSSGEKEQICLMWAYEKGHDAIVT--LLKHKXRPODELPCNEYSQPG 420  
 722 HOADYN-----AKTLGYSPFHQAQQGHTDIVTLLKXNGASP-----NEVSDG 766  
 421 GDG-----SYVSPSPGLGKIKMTKEKADIALL--RAGLPSHFH--LQISEIEFHEI 468  
 767 TPTPLAIAKRLGYISVTDVL---KVVTDTSFVLVSKHRMSPETVDEILDVSEDEGEEL 823  
 469 I 469  
 824 I 824  
 RESULT 3  
 ANKL\_MOUSE  
 ID ANKL\_MOUSE STANDARD; PRT; 1862 AA.  
 AC Q02357;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DT Ankyrin 1 (Erythrocyte ankyrin).  
 GN ANK1 OR ANK-1.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OC NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Erythrocyte;  
 RX MEDLINE=92345717; PubMed=1386265;  
 RA White R.A., Birkenmeier C.S., Peters L.L., Barker J.E., Lux S.E.;  
 RT "Murine erythrocyte ankyrin cDNA: highly conserved regions of the  
 RT regulatory domain."  
 RT Mamm Genome 3:281-285(1992).  
 RL  
 CC -!- FUNCTION: Attach integral membrane proteins to cytoskeletal  
 CC elements; bind to the erythrocyte membrane protein band 4.2, to  
 CC Na-K ATPase, to the lymphocyte membrane protein gp85, and to the  
 CC cytoskeletal proteins fodrin, tubulin, vimentin and desmin.  
 CC Erythrocyte ankyrins also link spectrin (beta chain) to the  
 CC cytoplasmic domain of the erythrocytes anion exchange protein;  
 CC they retain most or all of these binding functions.  
 CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC SURFACE OF ERYTHROCYTIC  
 CC PLASMA MEMBRANE

```
Query Match      10.5%; Score 460.5; DB 1; Length 1880;
Best Local Similarity 27.6%; Pred. No. 3.2e-20;
Matches 166; Conservative 76; Mismatches 176; Indels 183; Gaps 23;
```

CC -!- PTM: Regulated by phosphorylation (By similarity).  
 CC -!- PTM: Acylated by palmitic acid group(s) (By similarity).  
 CC -!- SIMILARITY: Contains 23 ANK repeats.  
 CC -!- SIMILARITY: Contains 1 death domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M84756; AAA37236.1; .  
 DR PIR; I49502; I49502.  
 DR HSSP; Q00420; 1AWC.  
 DR MGD; MGI:88024; Ankl.  
 DR InterPro; IPR002110; ANK.  
 DR InterPro; IPR000488; Death.  
 DR Pfam; PF00023; ank; 24.  
 DR Pfam; PF00531; death; 1.  
 DR Pfam; PF00791; ZUS; 1.  
 DR PRINTS; PRO1415; ANKYRIN.  
 DR SMART; SM00248; ANK; 22.  
 DR SMART; SM00005; DEATH; 1.  
 DR SMART; SM00218; ZUS; 1.  
 DR PROSITE; PS50088; ANK REPEAT; 20.  
 DR PROSITE; PS50297; ANK REP REGION; 1.  
 DR PROSITE; PS50017; DEATH DOMAIN; 1.  
 KW Cytoskeleton; Repeat; ANK repeat; Phosphorylation; Lipoprotein.  
 FT DOMAIN 1 827  
 FT 89 KDa DOMAIN (ANION EXCHANGE PROTEIN  
 FT BINDING DOMAIN).  
 FT 62 KDa DOMAIN (SPECTRIN BINDING  
 FT DOMAIN).  
 FT 55 KDa REGULATORY DOMAIN (REGULATES  
 FT THE BINDING OF ANKYRIN TO SPECTRIN  
 FT AND THE BAND 3 PROTEIN).  
 FT ANK 1.  
 FT ANK 2.  
 FT ANK 3.  
 FT ANK 4.  
 FT ANK 5.  
 FT ANK 6.  
 FT ANK 7.  
 FT ANK 8.  
 FT ANK 9.  
 FT ANK 10.  
 FT ANK 11.  
 FT ANK 12.  
 FT ANK 13.  
 FT ANK 14.  
 FT ANK 15.  
 FT ANK 16.  
 FT ANK 17.  
 FT ANK 18.  
 FT ANK 19.  
 FT ANK 20.  
 FT ANK 21.  
 FT ANK 22.  
 FT ANK 23.  
 FT DEATH.  
 FT DOMAIN 1399 1483  
 SQ SEQUENCE 1862 AA; 204242 MW; AE6B5B5B29001E5 CRC64;

Query Match 10.4%; Score 457.5; DB 1; Length 1862;  
 Best Local Similarity 28.7%; Pred. No. 4.9e-20;  
 Matches 151; Conservative 66; Mismatches 171; Indels 139; Gaps 17;  
 60 LNYATE-----NGLSLHLCCICGGKKSHRTIMLKLRLPSRLTRNGFTALFLAVYKDN 113  
 321 LQYNAEIDDTLDHLTLPLHVAHCGHHRV-AKVLLDKGAKPNSRALNGFTPLHIACKKNH 379  
 114 AELTSLHLSGADIQQVGYGGLTALHATAGHLEAADVLQGANVNCDAVFFFTPLHI 173

Db 380 IRVMELLKTKGASIDAVTESGLTPLHVASPMGHLPTVKNLQRGASPNVSNVKVEITPLHM 439  
 QY 174 AAYGHQVQVTRLLKFGADVNVSGEVDRLHLHLSAKAGFLNTAKLMESGSKADVNAQN 233  
 Db 440 AARAGHTEVAKYLLQNKAKANAKAKODQTPHCAARIGHTGMVKLLLENG--ASPNLAT 497  
 QY 234 EDHVPFLHFCRFGH-----HDIVKY-----LLQSD-- 258  
 Db 498 AGHTPLHTAREGHVDLTALLLKEASQACMTKGTPLHVAKYKVELAELLEHDAH 557  
 QY 259 -----LEV-----OPHVNIYGDTPPLHLACYNKFEVAKEI 289  
 Db 558 PNAAGKNGLTPLHVAHNNLDIVKLLPRGGSPHSPAWNGYTPLHIAAKQKQIEVARSL 617  
 QY 290 IQISGTESLTKENIFSETAFHSACTYCKSIDLVKFLLDQNV----- 330  
 Db 618 LQYCG--SANAESVQGVTPHLAAQSGHT-EMVALLLSQANGNLNKSGLTPLHLVQSE 674  
 QY 331 -----ININHQGRDGHGTGLHSACYHGHIRLVQFLDNGADMLNVACDPSSRS 377  
 Db 675 GHVLVADVLIKHGVTVDATTRMGYTPHLVASHYGNIKLVKFLQLHQADV-----A 725  
 QY 378 GKDEQTCIMWAYEKGHDAIVT-LLKHXYRQDELPCNEYSQPGDG-----SYVSVP 429  
 Db 726 KTKLGSPLHQAAGQGHDTIVTLLKNGASP-----NEVSSNGTTPPLAIAKRLGYISVT 779  
 QY 430 SPLGKIKSMTKERADIILL---RAGLPSPHF--LQLSEIEFHEIIGS 471  
 Db 780 DVL---KVVTDETSLVLSVDKHSYFETVDEILDVSEDEGDELVS 823  
 RESULT 4  
 ID\_KYK2 DICI STANDARD; PRT; 410 AA.  
 AC P18161;  
 DT 01-NOV-1990 (Rel. 16, Created)  
 DT 01-NOV-1990 (Rel. 16, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Tyrosine-protein kinase 2 (EC 2.7.1.112) (Fragment).  
 GN SPLB OR PYXB OR DPYK2.  
 OS Dictyostelium discoideum (Slime mold).  
 CC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.  
 OX NCBI\_TaxID=44689;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90287147; PubMed=1972546;  
 RA Tan J.L., Spudich J.A.;  
 RT "Developmentally regulated protein-tyrosine kinase genes in  
 RT Dictyostelium discoideum";  
 RL Mol. Cell. Biol. 10:3578-3583 (1990).  
 CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein  
 CC tyrosine phosphate.  
 CC -!- SIMILARITY: Belongs to the Tyr family of protein kinases.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; M33784; AAA33203.1; .  
 DR PIR; B35670; B35670.  
 DR HSSP; P08631; IADS.  
 DR DictyBase; DDB0002037; splB.  
 DR InterPro; IPR000719; Prot\_kinase.  
 DR InterPro; IPR001245; Tyr\_kinase.  
 DR Pfam; PF00069; pkinase; 1.  
 DR PRINTS; PR00109; TYRKINASE.  
 DR ProDom; PD000001; Prot\_kinase; 1.

DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.  
DR PROSITE; PS00109; PROTEIN KINASE TIR; 1.  
DR PROSITE; PS00111; PROTEIN KINASE DOM; 1.  
KW Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation.  
FT DOMAIN 108 381 PROTEIN KINASE.  
FT NP\_BIND 114 122 ATP (BY SIMILARITY).  
FT BINDING 135 135 ATP (BY SIMILARITY).  
FT ACT\_SITE 232 232 BY SIMILARITY.  
SQ SEQUENCE 410 AA; 46386 MW; E93918B6059AEC1 CRC64;  
  
Query Match 10.3%; Score 450.5; DB 1; Length 410;  
Best Local Similarity 31.1%; Pred. No. 1.8e-20;  
Matches 131; Conservative 80; Mismatches 145; Indels 65; Gaps 14;  
  
QY 365 DMNIVACDPSRSGSDEQCLMWAYEKGDHAIIVT---LLKHVYKQDDELPCNEYSPQGG 421  
DB 11 DITLVCDNDSTKEKNVS-----NTSSIIISASNLNRH-----ITNSHMRPG 55  
  
QY 422 DG-SYVSPGLGK-----IKMTKEKADILLRAGL-PSHPLQLSIEIPHEIIG 470  
DB 56 RSISGLIMSPINKESLNDIQRAIESEKIKTKPEELKSTILGEREYIIDINDIQTKVG 115  
  
QY 471 SGSPGVYKGRNKIIVAIKRYANTYCSKSDVDMFCREVSILCOLNHPVCFVGCALN 530  
DB 116 EGAFSEVWGWGKIHWAIKKLKI-IGDEQFERFIREVQNLKGNHONIVNFIGACYK 174  
  
QY 531 DPSQFAIVTYISGGSFSLHE-----QKRIILDOSKLIIVDVAKGMEYLNLTQPII 585  
DB 175 -PA-CIITEYMGAGSYLNILNPNSTPKVYSFPLVLKQADMGALGLHLHSIT--IV 229  
  
QY 586 HRDLNSHILLYEDHAVADFGESRFLQSLDEDNMTKQPG---NLRWAPVFTQCTRY 642  
DB 230 HRDLTSQNLIDELGNIKISDFGLS--AEKRSRGTMTWNGGICNPRWPPPELTKNLGHY 287  
  
QY 643 TIKADVSFALCLWEILTGPFAHLPAALAAADMAHYHRRPPIGYSIKPKISSLLIRGW 702  
DB 288 SEKYDVYCFSLVWEILTGPISDLDSQSORSQVAVAGLRPIPYCDPELKLITQCM 347  
  
QY 703 NACPEGRPEFSEVYMKLEBCLN-----IELMSPASNSGSLSP 742  
DB 348 EADPNDRPPFTYIVNKLKEISWNNPFIQVSDQFYQYSEPTPRLALNSQSSSISLSP 407  
  
QY 743 S 743  
DB 408 T 408

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
-----  
DR HSP; AB002377; BAA20833.2; -.  
DR HSP; P80144; 2MYO.  
DR InterPro; IPR002110; ANK.  
DR Pfam; PF00023; ank; 28.  
DR PRINTS; PR01415; ANKYRIN.  
DR SMART; SM00248; ANK; 28.  
DR PROSITE; PS50086; ANK\_REPEAT; 24.  
DR PROSITE; PS50297; ANK\_REPEAT\_REGION; 1.  
KW Hypothetical protein; Repeat; ANK repeat.  
FT NON\_TER 1 75 ANK 1.  
FT REPEAT 46 75 ANK 2.  
FT REPEAT 79 108 ANK 3.  
FT REPEAT 112 141 ANK 4.  
FT REPEAT 145 174 ANK 5.  
FT REPEAT 178 207 ANK 6.  
FT REPEAT 211 240 ANK 7.  
FT REPEAT 244 273 ANK 8.  
FT REPEAT 277 307 ANK 9.  
FT REPEAT 311 340 ANK 10.  
FT REPEAT 344 373 ANK 11.  
FT REPEAT 377 406 ANK 12.  
FT REPEAT 410 439 ANK 13.  
FT REPEAT 443 472 ANK 14.  
FT REPEAT 476 506 ANK 15.  
FT REPEAT 510 540 ANK 16.  
FT REPEAT 555 584 ANK 17.  
FT REPEAT 588 617 ANK 18.  
FT REPEAT 622 651 ANK 19.  
FT REPEAT 658 687 ANK 20.  
FT REPEAT 691 720 ANK 21.  
FT REPEAT 724 753 ANK 22.  
FT REPEAT 761 790 ANK 23.  
FT REPEAT 793 823 ANK 24.  
FT REPEAT 828 857 ANK 25.  
FT REPEAT 861 891 ANK 26.  
FT REPEAT 895 924 ANK 27.  
FT REPEAT 931 960 ANK 28.  
SQ SEQUENCE 1059 AA; 113465 MW; C1F55E6CFE494770 CRC64;  
  
Query Match 10.0%; Score 441; DB 1; Length 1059;  
Best Local Similarity 24.2%; Pred. No. 2.4e-19;  
Matches 214; Conservative 120; Mismatches 360; Indels 190; Gaps 40;  
  
QY 21 SESVITIERLEDDLOIKKEKELTELNRNIFGSD-----AFSKVNLNRYTENGSLHLHC 74  
DB 2 AEATAMAFKLIRQPSL-----VQAFNGDDPEYRALIFKEDVNFQDNEKRTPLHAA 54  
  
QY 75 CIOGKXSHIRITMLKGLRPSRLTRNGFTALHVAHYKDNAELITSLHSGADIQQVGYGG 134  
DB 55 AYLGDABEITELLISGARVNAKSKWLTPLHRAVASCSEEAQVQLLKHSAVDNARDKNW 113  
  
QY 135 LTALHIAITAGHLEAADVLHOGANNVNIQDAVFTPLHIAAYYGHQVTRLLIKFGADV 194  
DB 114 QTPHIAAANKAVKCAEALVPLLSNVNSDRAGRTALHRAAFSGHGMVKLLLSRGANIN 173  
  
QY 195 VSGEVGDRPLHLASAKGLNLIKLLMEEGSKADVNADNEDHVPPLHFCRFGHHIVKYL 254  
DB 174 AFDKDRRAIHWAAYMGHIEVVKLLVSHG--AEVCKDKKSYTFLHAAASSGMISVVKYL 231  
  
QY 255 LOSLEVQPHVNNIYGTPLHLACYNGKFEVAKETIQISGTSLTKEITFSETAFHSACT 314  
DB 232 L--DLGVDMPNPNAYGNTPHLVACYNGQDVVNNELIDCGAIVNQKNEKGTPLHFAAST 289  
  
QY 315 YGKSIDLVKFLLDQNVININHOGRDGHGTGLHSACYGHIRLVQFLDNGADMNLVACDPS 374



[illegible]

## RESULT 7

### ANK3 HUMAN

ID	ANK3_HUMAN	STANDARD;	PRT; 4377 AA.
AC	Q12955;		
AD	28-FEB-2003 (Rel. 41, Created)		
AE	28-FEB-2003 (Rel. 41, Last sequence update)		
AF	15-MAR-2004 (Rel. 43, Last annotation update)		
AG	Ankyrin 3 (ANK-3) (Ankyrin G).		
AH	ANK3.		
AI	Homo sapiens (Human).		
AJ	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AK	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AL	NCBI_TaxID=9606;		
AM	[1]		
AN	SEQUENCE FROM N.A.		
AO	TISSUE=Brain stem;		
AP	MEDLINE=95138209; PubMed=7836469;		
AQ	Kordeli E., Lambert S., Bennett V.;		
AR	"Ankyrin. A new ankyrin gene with neural-specific isoforms localized		
AS	at the axonal initial segment and node of Ranvier.";		
AT	J. Biol. Chem. 270:2352-2359 (1995).		
AV	-!- FUNCTION: Membrane-cytoskeleton linker. The neural-specific		
AW	isoforms may participate in the maintenance/targeting of ion		
AX	channels and cell adhesion molecules at the nodes of Ranvier and		
AY	axonal initial segments.		
AZ	-!- SUBUNIT: Neural-specific isoforms may be a constituent of a		
BA	neurofascin/NRCAM/ankyrin G complex.		
BB	-!- ALTERNATIVE PRODUCTS:		
BC	Event-Alternative splicing; Named isoforms=1;		
BD	Comment=A number of isoforms are produced;		
BE	Name=1; Synonyms=480-kDa isoform;		
BF	Isoid=Q12955-1; Sequences=Displayed;		
BG	-!- TISSUE SPECIFICITY: Expressed in brain and other tissues. Isoform		
BH	1 is neural-specific.		
BI	-!- SIMILARITY: Contains 23 ANK repeats.		
BJ	-!- SIMILARITY: Contains 1 death domain.		
BK	-----		
BL	This SWISS-PROT entry is copyright. It is produced through a collaboration		
BM	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
BN	the European Bioinformatics Institute. There are no restrictions on its		
BO	use by non-profit institutions as long as its content is in no way		
BP	modified and this statement is not removed. Usage by and for commercial		
BQ	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announcement/">http://www.isb-sib.ch/announcement/</a>		
BR	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).		
BS	-----		
BT	EMBL; U13616; AAA64834.1; -.		
BU	PIR; A55575; A55575.		
BV	HSP; P55273; 1B18.		
BW	Genew; HGNC:494; ANK3.		
BX	MIM; 600455; -.		
BY	GO; GO:0006605; P:protein targeting; NAS.		
BZ	InterPro; IPR002110; ANK.		
CA	InterPro; IPR000488; Death.		
CB	InterPro; IPR000906; ZU5.		
CC	Pfam; PF00023; ank; 24.		
CD	Pfam; PF00531; death; 1.		
CE	Pfam; PF00791; ZU5; 1.		
CF	PRINTS; PR01415; ANKYRIN.		
CG	SMART; SM00248; ANK; 21.		
CH	SMART; SM00005; DEATH; 1.		
CI	SMART; SM00218; ZU5; 1.		
CJ	PROSITE; PS50088; ANK_REPEAT; 21.		
CK	PROSITE; PS50297; ANK_REPEAT_REGION; 1.		
CL	PROSITE; PS50017; DEATH_DOMAIN; 1.		
CM	Cytoskeleton; Alternative splicing; Repeat; ANK repeat.		
CN	73 102 ANK 1.		
CO	FT REPEAT 106 135 ANK 2.		
CP	FT REPEAT 139 168 ANK 3.		
CQ	FT REPEAT 172 201 ANK 4.		
CR	FT REPEAT 203 230 ANK 5.		
CS	FT REPEAT 234 263 ANK 6.		
CT	FT REPEAT 267 296 ANK 7.		
CU	FT REPEAT 300 329 ANK 8.		
CV	FT REPEAT 333 362 ANK 9.		
CW	FT REPEAT 366 395 ANK 10.		



RESULT 9	
M3K7_HUMAN	
ID	M3K7_HUMAN
AC	043318; 043317; 043319; STANDARD; PRT; 506 AA.
DT	16-OCT-2001 (Rel. 40, Created)
DT	16-OCT-2001 (Rel. 40, Last sequence update)
DT	15-MAR-2004 (Rel. 43, Last annotation update)
DE	Mitogen-activated protein kinase kinase 7 (EC 2.7.1.1.)
DE	(transforming growth factor-beta-activated kinase 1) (TGF-beta-activated kinase 1).
GN	MAP3K7 OR TAK1.
OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX	NCBI_TaxID=9606;
RN	[1]
RC	SEQUENCE FROM N.A. (ISOFORMS 1A; 1B AND 1C).
RC	TISSUE=Lung;
RC	MEDLINE=22388257; PubMed=9480845;
RA	Sakurai H., Shigemori N., Hasegawa K., Sugita T.;
RT	"TGF-beta-activated kinase 1 stimulates NF-kappa B activation by an NF-kappa B-inducing kinase-independent mechanism.";
RL	Biochem. Biophys. Res. Commun. 243:545-549 (1998).
RN	[2]
RC	SEQUENCE FROM N.A. (ISOFORM 1A).
RC	TISSUE=Uterus;
RC	MEDLINE=22388257; PubMed=12477932;
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA	Hopkins R.F., Jordan H., Moore T., Max A.M., Wang J., Hsieh F.,
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA	Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,
RA	Baba S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA	Vallalath D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA	Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalish D.E.,
RA	Schneerch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT	"Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";
RT	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN	[3]
RC	INTERACTION WITH PELI1 AND PELI2.
RC	MEDLINE=22689054; PubMed=12804775;
RA	Jensen L.E., Whitehead A.S.;
RT	"Pellino2 activates the mitogen activated protein kinase pathway.";
RL	FEBS Lett. 545:199-202 (2003).
RN	[4]
RC	INTERACTION WITH PELI3.
RC	MEDLINE=22756745; PubMed=12874243;
RA	Jensen L.E., Whitehead A.S.;
RT	"Pellino3, a novel member of the Pellino protein family, promotes activation of c-Jun and Elk-1 and may act as a scaffolding protein.";
RL	J. Immunol. 171:1500-1506 (2003).
CC	-!- FUNCTION: Can phosphorylate and activate yet undefined MAPKs. Mediator of TGF-beta signal transduction. Stimulates NF-kappa-B activation.
CC	-!- SUBUNIT: Interacts with PELI1, PELI2 and PELI3.
CC	-!- ALTERNATIVE PRODUCTS:
CC	Event=Alternative splicing; Named isoforms=3;
CC	Name=1B;
CC	Isoid=043318-1; Sequence=Displayed;
CC	Name=1A;
CC	Isoid=043318-2; Sequence=VSP_004886;
CC	Name=1C;
CC	Isoid=043318-3; Sequence=VSP_004887; VSP_004888;
CC	-!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases. MAP

CC	kinase kinase subfamily.
CC	-!- SIMILARITY: IN THE C-TERMINAL SECTION; STRONG, TO C21ORF7.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation. The European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC	-----
CC	EMBL; ABC009357; BAA25026.1; -
DR	EMBL; ABC009356; BAA25025.1; -
DR	EMBL; ABC009358; BAA25027.2; -
DR	EMBL; BC017715; AAH17715.1; -
DR	PIR; JC5955; JC5955.
DR	PIR; JC5956; JC5956.
DR	HSP; P08631; IAD5.
DR	Genew; HGNC:16859; MAP3K7.
DR	MIM; 602614; -
DR	GO; GO:0004709; F:MAP kinase kinase activity; TAS.
DR	GO; GO:0007179; P:TGFbeta receptor signaling pathway; TAS.
DR	InterPro; IPR000719; Prot kinase.
DR	InterPro; IPR008271; Ser_thr_pkin_AS.
DR	InterPro; IPR001245; Tyr_pkinase.
DR	Pfam; PF00069; pkinase; 1.
DR	PRINTS; PR00109; TYRKINASE.
DR	ProDom; PD00001; Prot kinase; 1.
DR	PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR	PROSITE; PS00108; PROTEIN KINASE_ST; 1.
DR	PROSITE; PS00111; PROTEIN KINASE_DOM; 1.
KW	Transferase; Serine/threonine-protein kinase; ATP-binding;
KW	Alternative splicing.
FT	DOMAIN 8 14 POLY-SER.
FT	DOMAIN 36 291 PROTEIN KINASE.
FT	NP BIND 42 50 ATP (BY SIMILARITY).
FT	BINDING 63 63 ATP (BY SIMILARITY).
FT	ACT SITE 156 156 BY SIMILARITY.
FT	VARSPLIC 404 430 Missing (in isoform 1A).
FT	FTID=VSP_004886.
FT	VARSPPLIC 509 518 /PLAFCPNSKE -> ARTSCRTGPG (in isoform 1C).
FT	FTID=VSP_004887.
FT	VARSPPLIC 519 606 Missing (in isoform 1C).
FT	FTID=VSP_004888.
FT	VARSPPLIC 606 AA; 67196 MW; 3D8F8147CD174013 CRC64;
SEQ	SEQUENCE 606 AA; 67196 MW; 3D8F8147CD174013 CRC64;
	Query Match 9.5%; Score 419; DB 1; Length 606;
	Best Local Similarity 31.6%; Pred. No. 2.5e-18;
	Matches 115; Conservative 68; Mismatches 143; Indels 38; Gaps 13;
QY	458 LQLSETEFHEILGSGSGFYKGRCKNKIVATKRYANTYCSKSDVDMFCREVSLCOLN 517
DB	31 IDYKEIEVEEVGRGAFGVVCKAKRAKDVAKQTE-----SESEKAFIVELRQLSRVN 85
QY	518 HPCVIOFVGACLNDRSQFAIVTYQISGSLFSLH--EQKRLDLQSKLIIVADVAKGME 575
DB	86 HPMIVKLYGACLN--PVCVLMEYAEAGSLYVNLHGAEPYVYTAHAHSMWCLQCSQVA 142
QY	576 YLHNLTOP--IHRDLNSHLLIYEDCHAV-VADRGESRFLQSLDEDNMTQPGNLRWMA 632
DB	143 YLHSM-OPKALIHRLKPNLLVAGGTVKIKCDFGTACDIQT----HMTNKGGAAMWA 197
QY	633 PEVFTOCTRYTIKADVFSVALCLWEILTGEIIPFAHLKPAAAAADMAHYH-IRPFIYSIP 691
DB	198 PEVFE-EGSNVSEKCDVFSWGIILWEVITRRKFDGEGPAPRIMWAVHNGTRPLIKNLP 256
QY	692 KPTSSLLIRGNACPEGRPEFSEVNMKLECL-----CNIELMSA-----SSNSS 737
DB	257 KPTESLMTRCWSKDPQSRPSMEEIVIMTHLMRYFPFGADEPLQYPCQVSDGQSSATST 316
QY	738 GSLSPSSSDCLVNRGGPGRSHVAALRSRFELEYALNARSVAALSQAGYS---SQGLS 794
DB	317 GSFMDIASTN-TSKSDTNTMEQVPAINDIKRLSKLLKNQAKQSESGRLSLGASRGSS 375



FT	DOMAIN	135	141		POLY-GLY.
FT	DOMAIN	551	809		PROTEIN KINASE.
FT	NP_BIND	557	565		ATP (BY SIMILARITY).
FT	BINDING	578	578		ATP (BY SIMILARITY).
FT	ACT_SITE	676	676		BY SIMILARITY.
FT	MUTAGEN	596	596		E-S-K; IN CTR1-4; EXHIBITS ETHYLENE-
FT					TREATED PHENOTYPE.
FT	MUTAGEN	694	694		D-E; IN CTR1-1; EXHIBITS ETHYLENE-
FT					TREATED PHENOTYPE.
SQ	SEQUENCE	821 AA;	90306 MW;	2922D3DCDOCCI5BC CRC64;	

Query Match 9.2%; Score 404.5; DB 1; Length 821;  
Best Local Similarity 26.1%; Pred. No. 2.9e-17;  
Matches 139; Conservative 90; Mismatches 206; Indels 97; Gaps 20;

QY	273	PLHLACYNKFEFAKEIIQISGTE	:	:	:--SLTKENIFSFSTAHSACTYCKSID-	320
	:	:	:	:	:	:
DB	329	PMWKECIDGLKEIPFVVVPGISVLGRHRALLFKVLADIIDLPCRIAGCKVCYNRDDA	388			
QY	321	--LVKFLLDONVINHOQRDG-	---	TGLHSACYGHGIQLVQ-FLLDNGADMNL	368	
DB	399	ASCLVRFGLEDEYL-VDLVKPGHLMPEPSLLNGPSSISISSPLFRPKPVPAVFRL	447			
QY	369	VACDPSSRSSGEKDQTCMLMWAYBKGDHAIVTLKHVYRPQDELPNEYSQSGGGGSYV--	426			
	:	:	:	:	:	:
DB	448	LA-----KOYFSDSQS-LNLFVPDAPDDMGFSMFH-----RQYDNFPGGENDALAE	491			
QY	427	---SVSPGLGKIKSMTEKADILLRAGL-----PSHFHLQLS----	462			
DB	492	NCGGSLP-PSANPPQNMRASNQIEAAPWNAPPITQPVPNRANRELGLDGDMDI	550			
QY	463	IEPHEIIGSGFGKYKRCRNKIVAIKRYRANTYCYSKSDVDMPCREVSILCOLNHPCVI	522			
	:	:	:	:	:	:
DB	551	LNIRKEKIGAGSFGTVHRAEWHGSDVAVKILMEQDPAHER-VNEFLREVAIMKRLRHNPV	609			
QY	523	QPVGACLNDPSQFAIVTQYISGSLSLHQ--XAILDLQSKLIADVAKGYEYHLNL	580			
	:	:	:	:	:	:
DB	610	LFMGA-VTPPPNISIVTEYLSRGSLYELLHGSGAREQLDERRLSWAYDVAKGMNYLHNR	668			
QY	581	TOPIIHRLNSHNILLYEDGHAVVAOPFGESR-----FLOSLDEDNMTPQGNLRMAPEV	635			
	:	:	:	:	:	:
DB	669	NPPIVHRDLKSNLLVDKKYTVKVCDFGLSRUKASTFLSS-----KSAAGTPPWMAPEV	722			
QY	636	FQCCTRYTIKADVESYALCWEILTGEIIPAHIKPAAAAADYAHYHIRPPGTGYSTPKPIS	695			
	:	:	:	:	:	:
DB	723	LRD-EPSNEKSDVTSFGVILWELATLQCPWGNNLPAAVVAAGFYCKCLEIPRNLNPQVA	781			
QY	696	SILLRWNAACPGEPRSEFSEVVMKLECLCNIELMSPASSNSSGSSPSSSSD	747			
	:	:	:	:	:	:
DB	782	AIEGCWTNEFWKSPSATI-----MDLLRPLIKS----AVPPPNERSD	820			

RESULT 12  
HT16 HYDAT STANDARD; PRT; 757 AA.  
AC P5356;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Tyrosine-protein kinase Htk16 (EC 2.7.1.112).  
GN HTK16.  
OS Hydra attenuata (Hydra) (Hydra vulgaris).  
OC Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydroids; Anthomedusae;  
OC Hydridae; Hydra.  
OX NCBI\_TaxId=6087;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Irvine.  
RX MEDLINE=94181280; PubMed=8134129;  
RA Chan T.A.; Chu C.A.; Rauen K.A.; Kroher M.; Tatarewicz S.M.,  
RA Steele R.E.;  
RT "Identification of a gene encoding a novel protein-tyrosine kinase  
containing SH2 domains and ankryin-like repeats.";



RX MEDLINE=22388257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shermen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan B., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyx S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalusz D.E.,  
RA Schnerch A., Schein J.P., Jones S.J.M., Marra M.A.,  
RT "Generation and initial analysis of more than 15,000 full-length  
RT human and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [5]  
RN CHARACTERIZATION.  
RX MEDLINE=98409636; PubMed=9736715;  
RA Delcommenne M., Tan C., Gray V., Rue L., Woodgett J., Dedhar S.;  
RT "Phosphoinositide-3-OH kinase-dependent regulation of glycogen  
RT synthase kinase 3 and protein kinase B/AKT by the integrin-linked  
RT kinase.";  
RL Proc. Natl. Acad. Sci. U.S.A. 95:11211-11216(1998).  
RN [6]  
RN REVIEW  
RX MEDLINE=20179495; PubMed=10712922;  
RA Dedhar S.;  
RT "Cell-substrate interactions and signaling through ILK.";  
RL Curr. Opin. Cell Biol. 12:250-256(2000).  
CC -|- FUNCTION: RECEPTOR-PROXIMAL PROTEIN KINASE REGULATING INTEGRIN-  
CC MEDIATED SIGNAL TRANSDUCTION. MAY ACT AS A MEDIATOR OF INSIDE-OUT  
CC INTEGRIN SIGNALING. FOCAL ADHESION PROTEIN PART OF THE COMPLEX  
CC ILK-FINCH. THIS COMPLEX IS CONSIDERED TO BE ONE OF THE CONVERGENCE  
CC POINTS OF INTEGRIN- AND GROWTH FACTOR-SIGNALING PATHWAY. COULD BE  
CC IMPLICATED IN MEDIATING CELL ARCHITECTURE, ADHESION TO INTEGRIN  
CC SUBSTRATES AND ANCHORAGE-DEPENDENT GROWTH IN EPITHELIAL CELLS.  
CC PHOSPHORYLATES BETA-1 AND BETA-3 INTEGRIN SUBUNIT ON SERINE AND  
CC THREONINE RESIDUES, BUT ALSO AKT1 AND GSK3B.  
CC -|- ENZYME REGULATION: STIMULATED RAPIDLY BUT TRANSIENTLY BY BOTH CELL  
CC FIBRONECTIN INTERACTIONS, AS WELL AS BY INSULIN, IN A PI(3)K-  
CC DEPENDENT MANNER, LIKELY VIA THE BINDING OF PTDINS(3,4,5)P3 WITH A  
CC PH-LIKE DOMAIN OF ILK1.  
CC -|- SUBUNIT: INTERACTS WITH CYTOPLASMIC DOMAIN OF BETA 1 SUBUNIT OF  
CC INTEGRIN. COULD ALSO INTERACTS WITH BETA 2, BETA 3 AND/OR BETA 5  
CC SUBUNIT OF INTEGRIN. ALSO INTERACTS WITH PINCH AND PARVINS.  
CC -|- SUBCELLULAR LOCATION: Cytoplasmic.  
CC -|- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN HEART FOLLOWED BY SKELETAL  
CC MUSCLE, PANCREAS AND KIDNEY. WEAKLY EXPRESSED IN PLACENTA, LUNG  
CC AND LIVER.  
CC -|- DOMAIN: A PH-LIKE DOMAIN IS INVOLVED IN PHOSPHATIDYLINOSITOL  
CC PHOSPHATE BINDING.  
CC -|- PTM: Autophosphorylated on serine residues.  
CC -|- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.  
CC -|- SIMILARITY: Contains 3 ANK repeats.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: U40282; AAC16892.1; -;  
CC EMBL: AF244139; BAB74449.1; -;  
CC EMBL: AC404847; CAB99253.1; -;  
CC EMBL: BC001554; AAB01554.1; -;

DR HSP; Q00420; IAWC.  
DR Genew; HGNC:6040; ILK.  
DR MIM; 602366; -;  
DR GO; GO:0004574; F-protein serine/threonine kinase activity; TAS.  
DR GO; GO:0007160; P-cell-matrix adhesion; TAS.  
DR GO; GO:0007229; P-integrin-mediated signaling pathway; TAS.  
DR InterPro; IPR002110; ANK.  
DR InterPro; IPR000719; Prot kinase.  
DR InterPro; IPR0008271; Ser Thr pkin\_AS.  
DR InterPro; IPR001245; Tyr\_kinase.  
DR Pfam; PF00023; ank; 3.  
DR Pfam; PF00069; kinase; 1.  
DR PRINTS; PRO1415; ANKYRIN.  
DR PRINTS; PRO1019; TYRKINASE.  
DR SMART; SM00248; ANK; 3.  
DR ProDom; PD000001; Prot kinase; 1.  
DR PROSITE; PS50297; ANK\_REPEAT; 1.  
DR PROSITE; PS50088; ANK\_REPEAT; 3.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; FALSE\_NEG.  
DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; FALSE\_NEG.  
DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
DR Transferase; Serine/threonine-protein kinase; ATP-binding;  
KW Phosphorylation; Repeat; ANK repeat.  
FT REPEAT 33 82 ANK 1.  
FT REPEAT 66 95 ANK 2.  
FT REPEAT 99 128 ANK 3.  
FT DOMAIN 180 212 PH-LIKE.  
FT DOMAIN 193 446 PROTEIN\_KINASE.  
FT NP\_BIND 199 207 ATP (BY SIMILARITY).  
FT BINDING 220 220 ATP (BY SIMILARITY).  
FT MUTAGEN 359 359 E->K: INACTIVATION OF ILK.  
SQ SEQUENCE 452 AA; 51419 MW; E37DC2AD5311A1C2 CRC64;  
  
Query Match 9.1%; Score 400.5; DB 1; Length 452;  
Best Local Similarity 23.0%; Pred. No. 2.3e-17;  
Matches 128; Conservative 94; Mismatches 174; Indels 161; Gaps 15;  
  
Qy 192 DVNVSEVGDRLHLASAKGFLNIAKLMEBSKADVAQDNEDHVPFHFCRFGHHDIV 251  
Db 26 DLNQGDHGFSLHWACREGSAVVEMLMIRGA--INVMRGDDTPJHLAASHGHHRDV 83  
Qy 252 KYLLQSDLEVQHVNNVIGDTPLHLACVNGKEVAKEIIQISGTESLTAKENIFSETAFHS 311  
Db 84 QKLLQYKADI--NAVNEHGNVFLHYACFWGQDQVAEDLVA----- 121  
Qy 312 ACTYGSIDLVLKFLLDQNVININHOGRDHTGLHSACYHGHRLVQVFLLDNGADMLVAC 371  
Db 122 -----NGALVSI--- 128  
Qy 372 DPRSRSSEKDEQCLMWAVEKGHDAIVTLKHYPQDELPCNEYSQPGGGSYVSPSP 431  
Db 129 -----CNRYGEMPV-----KAKAP 143  
Qy 432 LGKIKSMTEKADILLR-----AGLPSHFHLQLSEIFHEIIGSGS 473  
Db 144 LRELLRERAEKMGONLRIPYKDTFWKGTTRPRNGTLNKGSGIDFKQLNFLTKLNEH 203  
Qy 474 FGKYYGRCRNKIVAKRYRANTYCKSDVDMFCRENSILQQLNHPCVIQFVGLCLNDPS 533  
Db 204 SGLWKGWRQGNDIVVKLVKVRDWSRTRSRD--FNEECPLRFIFSHPNVLPVIGACQSP 262  
Qy 534 QF-AIVTQVIGSGSLFSLHLEQKRLDLSKLI-IADVAKGMEYLNHLNLTQPIIHRD-LN 590  
Db 263 PPHTLTHMPYGSLYNVLHGTNFEVDSQAVKFDALDMARGVAFHLTL-EPLIPRHALN 321  
Qy 591 SHNLLYEDGHAVADFESRFLQSLDDEDMTKQ-PGNL-----RWMAPEVFTQCTRYTIK- 645  
Db 322 SRSMVDEDMTARI-----SNADVKFSFQCGRVPAPWAPVPEALQKPEDNRR 371  
Qy 646 -ADVFSYALCLWEILTGETPPRAHLKAPAAADMAADVAHHTRPPIGYSIPKPSILLIRGNA 704  
Db 372 SADMSFAVLLWELVETREVPFADLNMEIGMKVALEBURPTIPPGISPHVCKLMKICWNE 431



RA Janji B., Melchior C., Vallar L., Kieffer N.;  
RT "Cloning of an isoform of integrin-linked kinase (ILK) that is  
RT upregulated in HT-144 melanoma cells following TGF-beta1  
RT stimulation.";  
RL Oncogene 19:3069-3077(2000).  
CC -!- FUNCTION: RECEPTOR-PROXIMAL PROTEIN KINASE REGULATING INTEGRIN-  
CC MEDIATED SIGNAL TRANSDUCTION. MAY ACT AS A MEDIATOR OF INSIDE-OUT  
CC INTEGRIN SIGNALING. FOCAL ADHESION PROTEIN PART OF THE COMPLEX  
CC ILK-PINCH. THIS COMPLEX IS CONSIDERED TO BE ONE OF THE CONVERGENCE  
CC POINTS OF INTEGRIN- AND GROWTH FACTOR-SIGNALING PATHWAY. COULD BE  
CC IMPLICATED IN MEDIATING CELL ARCHITECTURE, ADHESION TO INTEGRIN  
CC SUBSTRATES AND ANCHORAGE-DEPENDENT GROWTH IN EPITHELIAL CELLS.  
CC PHOSPHORYLATES BETA-1 AND BETA-3 INTEGRIN SUBUNIT ON SERINE AND  
CC THREONINE RESIDUES, BUT ALSO AKT1 AND GSK3B (BY SIMILARITY).  
CC -!- SUBUNIT: INTERACTS WITH CYTOPLASMIC DOMAIN OF BETA 1 SUBUNIT OF  
CC INTEGRIN. COULD ALSO INTERACTS WITH BETA 2, BETA 3 AND/OR BETA 5  
CC SUBUNIT OF INTEGRIN. ALSO INTERACTS WITH PINCH AND PARVINS (BY  
CC similarity).  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
CC -!- TISSUE SPECIFICITY: Expressed in some highly invasive tumor cell  
CC lines but not in normal tissues.  
CC -!- INDUCTION: By TGF-beta1.  
CC -!- DOMAIN: A PH-LIKE DOMAIN IS INVOLVED IN PHOSPHATIDYLINOSITOL  
CC PHOSPHATE BINDING (BY SIMILARITY).  
CC -!- PTM: Autophosphorylated on serine residues.  
CC -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.  
CC -!- SIMILARITY: Contains 3 ANK repeats.  
CC -----  
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
CC EMBL; AJ277481; CAB94832.1; --  
CC HSP; Q00420; 1AWC.  
CC GO; GO:0005737; C:cytoplasm; NAS.  
CC GO; GO:0005524; F:ATP binding; NAS.  
CC GO; GO:0004674; F:protein serine/threonine kinase activity; IDA.  
CC GO; GO:0008283; P:cell proliferation; NAS.  
CC GO; GO:0007160; P:cell-matrix adhesion; NAS.  
CC GO; GO:0007229; P:integrin-mediated signaling pathway; NAS.  
CC GO; GO:0006468; P:protein amino acid phosphorylation; IDA.  
CC InterPro; IPR002110; ANK.  
CC InterPro; IPR000719; Prot\_kinase.  
CC InterPro; IPR008271; Ser\_thr\_pkin\_AS.  
CC Pfam; PF00023; ank; 3.  
CC Pfam; PF00069; pkinase; 1.  
CC PRINTS; PR01415; ANKYRIN.  
CC ProDom; PD000001; Prot\_kinase; 1.  
CC SMART; SM00248; ANK\_3\_kinase; 1.  
CC PROSITE; PS00297; ANK\_REPEAT; 1.  
CC PROSITE; PS00088; ANK\_REPEAT; 3.  
CC PROSITE; PS00107; PROTEIN\_KINASE\_ATP; FALSE\_NEG.  
CC PROSITE; PS00108; PROTEIN\_KINASE\_ST; FALSE\_NEG.  
CC PROSITE; PS00111; PROTEIN\_KINASE\_DOM; 1.  
CC Transferase; Serine/threonine-protein kinase; ATP-binding;  
CC Phosphorylation; Repeat; ANK repeat.  
CC REPEAT 33 62 ANK 1.  
CC REPEAT 66 95 ANK 2.  
CC REPEAT 99 128 ANK 3.  
CC FT DOMAIN 180 212 PH-LIKE.  
CC FT DOMAIN 193 446 PROTEIN\_KINASE.  
CC FT NP\_BIND 199 207 ATP (BY SIMILARITY).  
CC FT BINDING 220 220 ATP (BY SIMILARITY).  
CC SEQUENCE 452 AA; 51393 MW; BF520214E373DC88 CRC64;

Query Match 8.9%; Score 390.5; DB 1; Length 452;  
Best Local Similarity 22.8%; Pred. No. 9,5e-17;  
Matches 127; Conservative 94; Mismatches 175; Indels 161; Gaps 15;

QY 192 DVNVSGEVGDREPLHLASAKGFLNIAKLMEEGSKADVNAQDNEDHVLHFCSEFGHHDIV 251  
Db 26 DLNQGDDHGFSPLHWACREGRSAVVEMLMIEGAR--INVMRGDDTPLHLAASHGHREDIV 83  
QY 252 KYLLQSDLEVPQHVNYIGDTPHLHACVNGKFEVAKETIIQISGTESITKENIFSETAFHS 311  
Db 84 QKLLQYKADI--NAVNEHGNVPLHYACFWGQDQVAEDLVA----- 121  
QY 312 ACTYGKSIDLVKFLDDQNVININHQGRDHTGLHSACVHGHIRLVQVFLDNGADMNLVAC 371  
Db 122 -----NGALVSI--- 128  
QY 372 DPSRSSGEKDEQTCMLWAYEKGDHDAIVTLKHRYRPODELPCEYEQPGDGSVSVSP 431  
Db 129 -----CNKYGEMPDV-----KAKAP 143  
QY 432 LGKIKSMTEKADILLR-----AGLPSEHFLQLSETEFEHIIIGSGS 473  
Db 144 LRELLREAEKMGQNLNRIPYKDTFWKGTTRTPRNGTLNKGSGIDFKQLNFKLAKLNENH 203  
QY 474 FGKYGKGRCKNIVAIAKRYRANTYCSKSDVDMFCREVSILCOLNHPCVIOFVGAQLNDPS 533  
Db 204 SGELNKGSRWQSNDIVVKVLKVRDMSTRKSRD-FNEECPRLRIFGHPNVLPVLGACQAPPA 262  
QY 534 QF-AIVTOYISGSLFSLHSEOKRILDLSKLI-IAVDVAKGMEYLHNLTPQIIHRD-LN 590  
Db 263 PHPTLIITHMPYGSYLVNLHSGTNFVDQSQAVKFDLMARGMAFLHL-EPLIPRALN 321  
QY 591 SHNILLYEDGHAVVADFGESRFLOSLEDNNMTQ-PGNL---RWMapevftqctrytik- 645  
Db 322 SSSVMIDSDMTARI-----SMADVKEFQCPGRMVAPAMVAPEALQKKPFEDNRR 371  
QY 646 -ADVSYALCLWEILLTGIPFAHLKPAADAAADMAHYHRRPPIGYSIKPISILLIRGWA 704  
Db 372 SADMWSFALLWELVTREVPFADLSNMEIGMKVALEGLRPTIPPGISPHVCKLMKICMNE 431  
QY 705 CPEGRPEFSEVVMKLEE 721  
Db 432 DPAKRSKEDMIVPILK 448

Search completed: September 2, 2004, 14:21:46  
Job time : 39 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 5, 2004, 19:33:36 ; Search time 6441.43 Seconds  
(without alignments)  
16855.624 Million cell updates/sec

Title: US-10-626-173-3

Perfect score: 2505

Sequence: 1 aggggaattataatcttag.....gcagctttgaggaacagcgc 2505

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Capext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*

1: gb.ba.\*

2: gb.htg.\*

3: gb.in.\*

4: gb.om.\*

5: gb.ov.\*

6: gb.pat.\*

7: gb.ph.\*

8: gb.pl.\*

9: gb.pr.\*

10: gb.ro.\*

11: gb.sts.\*

12: gb.sy.\*

13: gb.un.\*

14: gb.vi.\*

15: em.ba.\*

16: em.fun.\*

17: em.hum.\*

18: em.in.\*

19: em.mu.\*

20: em.om.\*

21: em.or.\*

22: em.ov.\*

23: em.pat.\*

24: em.ph.\*

25: em.pl.\*

26: em.ro.\*

27: em.sts.\*

28: em.un.\*

29: em.vi.\*

30: em.htg.hum.\*

31: em.htg.inv.\*

32: em.htg.other.\*

33: em.htg.mus.\*

34: em.htg.pln.\*

35: em.htg.rtd.\*

36: em.htg.nam.\*

37: em.htg.vrt.\*

38: em.sy.\*

39: em.htgo.hum.\*

40: em.htgo.mus.\*

41: em.htgo.other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	2505	100.0	2505	6	BD269560	Novel CAR
2	2505	100.0	2505	6	BD269112	Sequence
3	2505	100.0	2505	6	AR437546	Sequence
4	2505	100.0	2508	6	AX056437	Sequence
5	2505	100.0	2508	9	AY303691	Homo sapi
6	2505	100.0	3001	9	AF116826	Homo sapi
7	2505	100.0	3025	6	BD269559	Novel CAR
8	2505	100.0	3025	6	AR437545	Sequence
9	2505	100.0	3025	6	AR437545	Sequence
10	1892.2	75.5	2505	6	BD269562	Novel CAR
11	1892.2	75.5	2505	6	AR269114	Sequence
12	1892.2	75.5	2505	6	AR437548	Sequence
13	1892.2	75.5	2505	6	AR437548	Sequence
14	1892.2	75.5	3026	10	AY303692	Novel CAR
15	1892.2	75.5	3026	6	BD269561	Sequence
16	1892.2	75.5	3026	6	AR437547	Sequence
17	1730.6	69.1	7507	9	HSN807040	Homo sapi
18	166.4	18.6	568	6	AX525654	Sequence
19	199.4	8.0	150272	2	AC073887	Homo sapi
20	199.4	8.0	179582	9	AC093158	Homo sapi
21	196.2	7.8	66066	2	AC026566	Homo sapi
22	177.6	7.1	112358	9	AC105271	Homo sapi
23	166.6	6.7	197553	10	AC124174	Mus muscu
24	165	6.6	269035	2	AC126082	Rattus no
25	185	6.6	276946	2	AC115396	Rattus no
26	153.4	6.1	150272	2	AC073887	Homo sapi
27	153	6.1	146018	2	AC018404	Homo sapi
28	149.4	6.0	66066	2	AC026566	Homo sapi
29	138.2	5.5	258378	2	AC105566	Rattus no
30	110.4	4.4	78331	9	AC098692	Homo sapi
31	107.2	4.3	60037	2	AC100314	Mus muscu
32	107.2	4.3	254819	2	AC144762	Mus muscu
33	106.4	4.2	1672	8	AY096470	Arabidops
34	106.4	4.2	2122	8	AY070086	Arabidops
35	101.4	4.0	28196	9	AC119672	Homo sapi
36	100.2	4.0	1704	6	AX652967	Sequence
37	100.2	4.0	2263	8	AK072690	Oryza sat
38	99.4	4.0	1748	8	AY128938	Arabidops
39	99.4	4.0	2104	8	AY093017	Arabidops
40	98	3.9	1662	6	AX506319	Sequence
41	94.8	3.8	1961	8	AK073747	Oryza sat
42	94	3.8	597	12	AY195855	Synthetic
43	92.6	3.7	1006	8	AK060220	Oryza sat
44	92.6	3.7	2251	8	AK103704	Oryza sat
45	91.4	3.6	3323	6	AX746788	Sequence

# ALIGNMENTS

RESULT 1  
BD269560  
LOCUS BD269560  
DEFINITION Novel CAR protein and nucleic acid molecules and uses therefor.  
ACCESSION BD269560  
VERSION BD269560.1 GI:33079328  
KEYWORDS JP 2002535962-A/2.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 2505)  
AUTHORS Raju, J.  
TITLE Novel CAR protein and nucleic acid molecules and uses therefor  
JOURNAL Patent: JP 2002535962-A 2 29-OCT-2002;

COMMENT	MILLENNIUM PHARMACEUTICALS INC OS Homo sapiens (human) PN JP 2002535962-A/2 PD 29-OCT-2002 PF 10-DEC-1999 JP 2000586772 PR 11-DEC-1998 US 60/111938,14-APR-1999 US 09/291839 PI JEYASEELAN RAJU PC C12N15/09,C07K16/40,C12N1/15,C12N1/19,C12N1/21,C12N5/10,C12N9/10,C12Q1/02, PC C12Q1/68,G01N33/15,G01N33/50,G01N33/53,G01N33/68//A61K38/45, PC A61K39/395 PC A61K45/00,A61P9/04,A61P9/00,A61P35/00,A61P43/00,C12N15/00,PC C12N5/00, PC A61K37/52 CC Novel CARP protein and nucleic acid molecules and uses CC therefor
FEATURES	key Location/Qualifiers FT CDS Location/Qualifiers 1..2505 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606"
ORIGIN	Query Match 100.0%; Score 2505; DB 6; Length 2505; Best Local Similarity 100.0%; Pred. No. 0; Matches 2505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  QY 1 ATGGGAAATTTAAATCTAGACCAACCCAAACTTGTACTGTAGTGAATGGAAGAAAAAGTC 60 DB 1 ATGGGAAATTTAAATCTAGACCAACCCAAACTTGTACTGTAGTGAATGGAAGAAAAAGTC 60  QY 61 AGTGAAATCATATGTTATCACAAATAGAAAAGATTAGAAGATGACCTGCAGATCAAGAAAAA 120 DB 61 AGTGAATCATATGTTATCACAAATAGAAAAGATTAGAAGATGACCTGCAGATCAAGAAAAA 120  QY 121 GAATGACAGAACTAAGGAATATATTGGCTCTGTATGAAGCCCTTCAGTAAAGTCAATTTA 180 DB 121 GAATGACAGAACTAAGGAATATATTGGCTCTGTATGAAGCCCTTCAGTAAAGTCAATTTA 180  QY 181 AATTACCGCACTGAAAATGGCTGTCTCTACTTCATTTAATGTTGCAATTTGTGGAGCAAG 240 DB 181 AATTACCGCACTGAAAATGGCTGTCTCTACTTCATTTAATGTTGCAATTTGTGGAGCAAG 240  QY 241 AATATCATATTCGAACTCTTATGTTGAAGGCTCCGCCATCTCGACTGACAGAATAAT 300 DB 241 AATATCATATTCGAACTCTTATGTTGAAGGCTCCGCCATCTCGACTGACAGAATAAT 300  QY 301 GGATTTACAGCCTTGCATTTAGCAGTTTACAAGGATAATGCAGAAATTCATCTCTCTG 360 DB 301 GGATTTACAGCCTTGCATTTAGCAGTTTACAAGGATAATGCAGAAATTCATCTCTCTG 360  QY 361 CTTACAGTGGAGCTGATATACAGCAGTTTGGATACGGTGGCTCACTGCCCTCCATATT 420 DB 361 CTTACAGTGGAGCTGATATACAGCAGTTTGGATACGGTGGCTCACTGCCCTCCATATT 420  QY 421 GCTACAATAGCTGGCCACCTAGAGGCTGTGTGTCTGTGTGCAACATGGAGCTAATGTC 480 DB 421 GCTACAATAGCTGGCCACCTAGAGGCTGTGTGTCTGTGTGCAACATGGAGCTAATGTC 480  QY 481 AATATTCAGATGAGAGTTTTCCTACATTCATATTTGACGGTACTATGGACATGAA 540 DB 481 AATATTCAGATGAGAGTTTTCCTACATTCATATTTGACGGTACTATGGACATGAA 540  QY 541 CAGTAACTCGCCTCTTTTGAATTTGGTGTGTATGTAATGTAAGTGGTGAAGTTGGA 600 DB 541 CAGTAACTCGCCTCTTTTGAATTTGGTGTGTATGTAATGTAAGTGGTGAAGTTGGA 600  QY 601 GATAGACCCCTCCACCTAGCAATCTGAAAAGAGATTTTGAATTTGCAAACTTTGATG 660 DB 601 GATAGACCCCTCCACCTAGCAATCTGAAAAGAGATTTTGAATTTGCAAACTTTGATG 660

Db 1741 ACAGCCAAATATACATCGTACCTTGACAGCTCACATATTTCTCTATGAGATGGG 1800  
Qy 1801 CATGCTGTGTGGCAGATTTTGAGAAATCAAGATTTCTACAGTCTCTGGATGAGACAAC 1860  
Db 1801 CATGCTGTGTGGCAGATTTTGAGAAATCAAGATTTCTACAGTCTCTGGATGAGACAAC 1860  
Qy 1861 ATGACAAAACAACTCGGAACTCCGTTGGATGGCTCCTGAGGTGTTCCAGCGAGTGCACT 1920  
Db 1861 ATGACAAAACAACTCGGAACTCCGTTGGATGGCTCCTGAGGTGTTCCAGCGAGTGCACT 1920  
Qy 1921 CGGTACACCATCAAAACAGATGTTCTACGTATGCTCTGTGTCTGTGGAAATTTCTCACT 1980  
Db 1921 CGGTACACCATCAAAACAGATGTTCTACGTATGCTCTGTGTCTGTGGAAATTTCTCACT 1980  
Qy 1981 GCGGAATTTCCATTCCTCATCTCAAGCAGCGGCTGGGAGCAGAGATGGCTTACCAC 2040  
Db 1981 GCGGAATTTCCATTCCTCATCTCAAGCAGCGGCTGGGAGCAGAGATGGCTTACCAC 2040  
Qy 2041 CACATCAGACCTCCCATTTGGCTATTCATTCCTCAAGCCCATATCATCTCTGCTGATCGA 2100  
Db 2041 CACATCAGACCTCCCATTTGGCTATTCATTCCTCAAGCCCATATCATCTCTGCTGATCGA 2100  
Qy 2101 GGGTGAAGCGCATGCTCTGAAGAGACCCGAATTTCTGAAGTTGTCTGAAGTTAGAA 2160  
Db 2101 GGGTGAAGCGCATGCTCTGAAGAGACCCGAATTTCTGAAGTTGTCTGAAGTTAGAA 2160  
Qy 2161 GAGTGTCTCTGAAACATTGAGTGTCTCTGATGATCAAGTAAACAGCAGTGGGTCTCTC 2220  
Db 2161 GAGTGTCTCTGAAACATTGAGTGTCTCTGATGATCAAGTAAACAGCAGTGGGTCTCTC 2220  
Qy 2221 TCACCTTCTCTCTCTGATTCGCTGTGTAACCGGAGGACCTGGCCGAGTCATGTG 2280  
Db 2221 TCACCTTCTCTCTCTGATTCGCTGTGTAACCGGAGGACCTGGCCGAGTCATGTG 2280  
Qy 2281 GCAGCATTAAGAGTGGTTTCGAATTTGGAATGATGCTCTAAGTCTGCTCTGAGTCTGCT 2340  
Db 2281 GCAGCATTAAGAGTGGTTTCGAATTTGGAATGATGCTCTAAGTCTGCTCTGAGTCTGCT 2340  
Qy 2341 TTGTCGCCAAAGTCTGAGCAATATTCCTCTCAAGTCTGCTCTTTGGAGGAGATGAAAAGA 2400  
Db 2341 TTGTCGCCAAAGTCTGAGCAATATTCCTCTCAAGTCTGCTCTTTGGAGGAGATGAAAAGA 2400  
Qy 2401 AGTCTTCAATACACCCATTGACAAATATGCTATGATTCGATCCCATGAGTCTCATG 2460  
Db 2401 AGTCTTCAATACACCCATTGACAAATATGCTATGATTCGATCCCATGAGTCTCATG 2460  
Qy 2461 CATTTTCATTCTTGCCGAAATAGTAGCAGCTTTGAGGACAGCAGC 2505  
Db 2461 CATTTTCATTCTTGCCGAAATAGTAGCAGCTTTGAGGACAGCAGC 2505

RESULT 2  
AR269112  
LOCUS AR269112 2505 bp DNA linear PAT 10-APR-2003  
DEFINITION Sequence 3 from patent US 6500654.  
ACCESSION AR269112  
VERSION AR269112.1 GI:29699955  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 2505)  
AUTHORS Raju,J.  
TITLE CARP protein and nucleic acid molecules and uses therefor  
JOURNAL Patent: US 6500654-A 3 31-DEC-2002;  
FEATURES Location/Qualifiers  
source 1..2505  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN  
Query Match 100.0%; Score 2505; DB 6; Length 2505;

Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGGGAAATTTATAATCTAGACCAACCCAAACTTTGTACTGTATGAATGGAAGAAAAGTC 60  
Db 1 ATGGGAAATTTATAAATCTAGACCAACCCAAACTTTGTACTGTATGAATGGAAGAAAAGTC 60  
Qy 61 AGTGAATCATATGTTTATCACAAATGAAAGATTAGAAAGTACCTCGAGATCAAGGAAAA 120  
Db 61 AGTGAATCATATGTTTATCACAAATGAAAGATTAGAAAGTACCTCGAGATCAAGGAAAA 120  
Qy 121 GAACTGACAGAACTAAGGAATATATTTGGCTCTGATGAAGCCTTCAGTAAAGTCAATTTA 180  
Db 121 GAACTGACAGAACTAAGGAATATATTTGGCTCTGATGAAGCCTTCAGTAAAGTCAATTTA 180  
Qy 181 AATTACCGCAGCTGAAATGCGGCTGCTCTACTCTATTTATTTGTCATTTGTGGAGGCAAG 240  
Db 181 AATTACCGCAGCTGAAATGCGGCTGCTCTACTCTATTTATTTGTCATTTGTGGAGGCAAG 240  
Qy 241 AAATCACATATTCGAACCTCTTATTTGAAAGGGCTCCGCCATCTCGACTGACAAGAAAT 300  
Db 241 AAATCACATATTCGAACCTCTTATTTGAAAGGGCTCCGCCATCTCGACTGACAAGAAAT 300  
Qy 301 GGATTTACAGCCTTGCAATTTAGCAGTTTACAAGGATATGAGAAATGATCATTCTCTG 360  
Db 301 GGATTTACAGCCTTGCAATTTAGCAGTTTACAAGGATATGAGAAATGATCATTCTCTG 360  
Qy 361 CTTTCAGTGGAGCTGATATACAGCAGGTTGGATACGGTGGCTCCTCACTGCCCTCCATATT 420  
Db 361 CTTTCAGTGGAGCTGATATACAGCAGGTTGGATACGGTGGCTCCTCACTGCCCTCCATATT 420  
Qy 421 GGTACAATAGCTGGCCACCTAGAGGCTGCTGATGCTGTTGCAACATGAGAGCTAAATGTC 480  
Db 421 GGTACAATAGCTGGCCACCTAGAGGCTGCTGATGCTGTTGCAACATGAGAGCTAAATGTC 480  
Qy 481 AATATTCAGATGCGAGTTTCTTCTCACTCCATTCGATATGAGGCTGCTATGGACATGAA 540  
Db 481 AATATTCAGATGCGAGTTTCTTCTCACTCCATTCGATATGAGGCTGCTATGGACATGAA 540  
Qy 541 CAGGTAACCTCGCTTCTTTTGAATTTGGTGTCTGATGTAATGTAAGTGGTGAAGTTGGA 600  
Db 541 CAGGTAACCTCGCTTCTTTTGAATTTGGTGTCTGATGTAATGTAAGTGGTGAAGTTGGA 600  
Qy 601 GATAGACCCCTCCACTAGCATCTGCAAAAGGATTTCTGAATATGCAAACTCTTGATG 660  
Db 601 GATAGACCCCTCCACTAGCATCTGCAAAAGGATTTCTGAATATGCAAACTCTTGATG 660  
Qy 661 GAAGAGGCGAGCAAGCAGATGTAATGCTCAAGATAATGAAGACCATGTCCCACTCCAT 720  
Db 661 GAAGAGGCGAGCAAGCAGATGTAATGCTCAAGATAATGAAGACCATGTCCCACTCCAT 720  
Qy 721 TTCTGTTCTCGATTTGGACACCATGATATAGTTAAGTATCTGCTGCAAAAGTGAATGGAA 780  
Db 721 TTCTGTTCTCGATTTGGACACCATGATATAGTTAAGTATCTGCTGCAAAAGTGAATGGAA 780  
Qy 781 GTTCAACCTCATGTTGTTAAATATCTATGAGATACCCCTTACACCTGGCATGCTACAT 840  
Db 781 GTTCAACCTCATGTTGTTAAATATCTATGAGATACCCCTTACACCTGGCATGCTACAT 840  
Qy 841 GGCAAATTTGAAAGTTGCCAAGGAAATCATCCAAATATCAGGAACAGAAAGTCTGACTAAG 900  
Db 841 GGCAAATTTGAAAGTTGCCAAGGAAATCATCCAAATATCAGGAACAGAAAGTCTGACTAAG 900  
Qy 901 GAAACATCTTCAGTGAACACAGCTTTTCATAGTGTGTTCTATGTCGCAAGCATTTGAC 960  
Db 901 GAAACATCTTCAGTGAACACAGCTTTTCATAGTGTGTTCTATGTCGCAAGCATTTGAC 960  
Qy 961 CTAGTCAAAATTTCTTCTTGATCAGAAATGTCATAAACATCAACCAAGGAGGATGGG 1020  
Db 961 CTAGTCAAAATTTCTTCTTGATCAGAAATGTCATAAACATCAACCAAGGAGGATGGG 1020  
Qy 1021 CACACTGGATTAACCTCTGCTGTGTACCAAGGTCATTCGGCTGGTTCACTTCTTACTG 1080  
Db 1021 CACACTGGATTAACCTCTGCTGTGTACCAAGGTCATTCGGCTGGTTCACTTCTTACTG 1080

Db 1021 CACACTGGATTACACTCTGCTTGTCTGCTACCAACGGTCACATTCGCCCTGCTTTCAGTTCTTACTG 1080  
Qy 1081 GATTAATGGAGCTGATATGAATCTAGTGGCTTTGTGATCCAGCAGAGTCTAGTGGTGAAGAAA 1140  
Db 1081 GATTAATGGAGCTGATATGAATCTAGTGGCTTTGTGATCCAGCAGAGTCTAGTGGTGAAGAAA 1140  
Qy 1141 GATCAGCAGACATGTTTGTGATGTTGAGTTATGAAAAAGGCGATGATCCATTTGTCACTC 1200  
Db 1141 GATCAGCAGACATGTTTGTGATGTTGAGTTATGAAAAAGGCGATGATCCATTTGTCACTC 1200  
Qy 1201 CTGAAGCAATTATAAGAGACCACAGATGAATTCCTGTAATGAATATTTCTCAGCTGGA 1260  
Db 1201 CTGAAGCAATTATAAGAGACCACAGATGAATTCCTGTAATGAATATTTCTCAGCTGGA 1260  
Qy 1261 GGAGATGGCTCTATGTTGTTGTTCCATCACCTTGGGGAAGATTAAAGAGATGACAAA 1320  
Db 1261 GGAGATGGCTCTATGTTGTTGTTCCATCACCTTGGGGAAGATTAAAGAGATGACAAA 1320  
Qy 1321 GAGAAGGCAGATATTTCTCCTTAAGAGCTGATTCCTTTCACATTTCCATTTCCAGTCT 1380  
Db 1321 GAGAAGGCAGATATTTCTCCTTAAGAGCTGATTCCTTTCACATTTCCATTTCCAGTCT 1380  
Qy 1381 TCAGAAATGAGTTCCATGAGATATTTGGCTCAGGTTCTTTTGGGAAGTATATAAGGA 1440  
Db 1381 TCAGAAATGAGTTCCATGAGATATTTGGCTCAGGTTCTTTTGGGAAGTATATAAGGA 1440  
Qy 1441 CGATGCAGAAATAAATAGTGGCTATATAAACGTTATCGAGCCAAATACCTACTGCTCCAAG 1500  
Db 1441 CGATGCAGAAATAAATAGTGGCTATATAAACGTTATCGAGCCAAATACCTACTGCTCCAAG 1500  
Qy 1501 TCAGATGGATGATGTTTGGCGAGAGGTGTCATTTCTCTGCCAGTCAATCATCCCTGC 1560  
Db 1501 TCAGATGGATGATGTTTGGCGAGAGGTGTCATTTCTCTGCCAGTCAATCATCCCTGC 1560  
Qy 1561 GTAATTCAGTTTGGGGTGTCTGTTGAATGATCCAGCCAGTTTCCCAATTTGCTCACTCA 1620  
Db 1561 GTAATTCAGTTTGGGGTGTCTGTTGAATGATCCAGCCAGTTTCCCAATTTGCTCACTCA 1620  
Qy 1621 TACATATCAGGGGTTCTCTGTTCTCCCTTCATGAGCAGAGAGATTTCTGATTTG 1680  
Db 1621 TACATATCAGGGGTTCTCTGTTCTCCCTTCATGAGCAGAGAGATTTCTGATTTG 1680  
Qy 1681 CAGTCTAAATTAATATTTGAGTAGATGTTGCCAAGGATGGAGTACCTTCACACCTG 1740  
Db 1681 CAGTCTAAATTAATATTTGAGTAGATGTTGCCAAGGATGGAGTACCTTCACACCTG 1740  
Qy 1741 ACACAGCCAAATATACATCGTGACTTGAACAGTCACAAATATTTCTCTATGAGGATGG 1800  
Db 1741 ACACAGCCAAATATACATCGTGACTTGAACAGTCACAAATATTTCTCTATGAGGATGG 1800  
Qy 1801 CATGCTGTGGGAGATTTTGGAGAAATCAAGATTTCTACAGTCTCTGATGAAGACAA 1860  
Db 1801 CATGCTGTGGGAGATTTTGGAGAAATCAAGATTTCTACAGTCTCTGATGAAGACAA 1860  
Qy 1861 ATGACAAAACAACTGGGAACCTCCGTTGGATGGCTCCTGAGTGTTCACGAGTGCACT 1920  
Db 1861 ATGACAAAACAACTGGGAACCTCCGTTGGATGGCTCCTGAGTGTTCACGAGTGCACT 1920  
Qy 1921 CGGTACACCAATCAAGCAGATGTTTTCAGTATGCTCTGTTGTTGGGAAATTTCTACT 1980  
Db 1921 CGGTACACCAATCAAGCAGATGTTTTCAGTATGCTCTGTTGTTGGGAAATTTCTACT 1980  
Qy 1981 GGGGAAATTCATTTCCATTCATCTCAAGCAGCGCTGCGCAGCAGACATGCTTACCAC 2040  
Db 1981 GGGGAAATTCATTTCCATTCATCTCAAGCAGCGCTGCGCAGCAGACATGCTTACCAC 2040  
Qy 2041 CACATCAGACCTCCCATTTGGCTATTTCCATTTCCCAAGCCCAATATCATCTCTGTGATAGA 2100  
Db 2041 CACATCAGACCTCCCATTTGGCTATTTCCATTTCCCAAGCCCAATATCATCTCTGTGATAGA 2100  
Qy 2101 GGGTGGAAACGATGTTCTGAGGAAGACCCGAAATTTTCTGAGTCTGATGAGTTAGAA 2160

Qy 2161 GAGTGTCTGTGCAACATTGAGCTGATGTCTCTGTCATCAAGTAACAGCAGTGGGTCTCTC 2220  
Db 2161 GAGTGTCTGTGCAACATTGAGCTGATGTCTCTGTCATCAAGTAACAGCAGTGGGTCTCTC 2220  
Qy 2221 TCACCTTCTTCTTCTTCTGATTCCTCTGATGAAACCGGGAGGACCTGGCCGAGTCATGTG 2280  
Db 2221 TCACCTTCTTCTTCTTCTGATTCCTCTGATGAAACCGGGAGGACCTGGCCGAGTCATGTG 2280  
Qy 2281 GCAGCATTAAGAGTGGTTTCGAATTGGAATATGCTCTAAATGCAAGTCTCTATGCTGCT 2340  
Db 2281 GCAGCATTAAGAGTGGTTTCGAATTGGAATATGCTCTAAATGCAAGTCTCTATGCTGCT 2340  
Qy 2341 TTGTCCCAAGTGTGCGACAATATTTCTCTCAAGTCTGTCTTTGGAGGAGATGAAAAGA 2400  
Db 2341 TTGTCCCAAGTGTGCGACAATATTTCTCTCAAGTCTGTCTTTGGAGGAGATGAAAAGA 2400  
Qy 2401 AGTCTTCAATACACACCCATTCACAAATATGCTATGATCCGATCCCATGAGCTCAATG 2460  
Db 2401 AGTCTTCAATACACACCCATTCACAAATATGCTATGATCCGATCCCATGAGCTCAATG 2460  
Qy 2461 CATTTTCATTTGCGAAATAGTAGCAGCTTTGAGGACAGCAGC 2505  
Db 2461 CATTTTCATTTGCGAAATAGTAGCAGCTTTGAGGACAGCAGC 2505

RESULT 3  
AR437546  
LOCUS AR437546 2505 bp DNA linear PAT 18-DEC-2003  
DEFINITION Sequence 3 from patent US 6660490.  
ACCESSION AR437546  
VERSION AR437546.1 GI:40202641  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 2505)  
AUTHORS Raju,J.  
TITLE CARK protein and nucleic acid molecules and uses therefor  
JOURNAL Patent: US 6660490-A 3 09-DEC-2003;  
FEATURES  
source Location/Qualifiers  
1..2505  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN  
Query Match 100.0%; Score 2505; DB 6; Length 2505;  
Best local Similarity 100.0%; Pred. No. 0;  
Matches 2505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGGGAAATTATAAATCTAGACCAACCCAAACTTGTACTGTAGTGAATGGAAGAAAAAGTC 60  
Db 1 ATGGGAAATTATAAATCTAGACCAACCCAAACTTGTACTGTAGTGAATGGAAGAAAAAGTC 60  
Qy 61 AGTGAATCATATGTTATCATCAATAGAAAGATTAGAAGATGACCTGCAGATCAAGAAAAA 120  
Db 61 AGTGAATCATATGTTATCATCAATAGAAAGATTAGAAGATGACCTGCAGATCAAGAAAAA 120  
Qy 121 GAATCAGACAGACTAGGAATATATTTGCTCTGATGAAGCTTCAGTAAAGTCAATTTA 180  
Db 121 GAATCAGACAGACTAGGAATATATTTGCTCTGATGAAGCTTCAGTAAAGTCAATTTA 180  
Qy 181 AATTACCGCACTGAAAATGGGCTGTCTCTACTTCTATTTATGTTGCAATTTGTGGAGGCAAG 240  
Db 181 AATTACCGCACTGAAAATGGGCTGTCTCTACTTCTATTTATGTTGCAATTTGTGGAGGCAAG 240  
Qy 241 AATATCATATTCGAATCTTTATGTTGAAGGGCTCCGCCATCTCGACTCACAGAAAT 300  
Db 241 AATATCATATTCGAATCTTTATGTTGAAGGGCTCCGCCATCTCGACTCACAGAAAT 300  
Qy 301 GGATTTACAGCCTTGTCAATTTAGCAGTTTACAAGGATAATGCAAGATTGATCACTTCTCTG 360  
Db 301 GGATTTACAGCCTTGTCAATTTAGCAGTTTACAAGGATAATGCAAGATTGATCACTTCTCTG 360

361 QY CTTACAGTGGAGCTGATATACAGCAGGTTGGATACGGTGGCCCTCACTGCCCTCCATATT 420  
361 Db CTTACAGTGGAGCTGATATACAGCAGGTTGGATACGGTGGCCCTCACTGCCCTCCATATT 420  
421 QY GCTACAAATAGCTGGCCACCTAGAGGCTGCTGATGCTGTTGGCAACATGAGAGCTAAATGTC 480  
421 Db GCTACAAATAGCTGGCCACCTAGAGGCTGCTGATGCTGTTGGCAACATGAGAGCTAAATGTC 480  
481 QY AATATTCAAGATGCAATGCTGTTTCTCACTCCATGCAATATGAGCGTACTATGGACATGAA 540  
481 Db AATATTCAAGATGCAATGCTGTTTCTCACTCCATGCAATATGAGCGTACTATGGACATGAA 540  
541 QY CAGGTAACCTGCCCTCTTTTGAATTTGGTGTGATGTAATGTAAGTGGTGAAGTTGGA 600  
541 Db CAGGTAACCTGCCCTCTTTTGAATTTGGTGTGATGTAATGTAAGTGGTGAAGTTGGA 600  
601 QY GATAGACCCCTCCACCTAGCATCTGCAAAAGGATTTCTGAAATATTGCAAACTCTTGATG 660  
601 Db GATAGACCCCTCCACCTAGCATCTGCAAAAGGATTTCTGAAATATTGCAAACTCTTGATG 660  
661 QY GAAGAAGGCGCAAAAGCAGATGTAATGCTCAAGATAATGAAGACCATGTGCCACTCCAT 720  
661 Db GAAGAAGGCGCAAAAGCAGATGTAATGCTCAAGATAATGAAGACCATGTGCCACTCCAT 720  
721 QY TTCTGTTCTGATTTGGACACCATGATATAGTTAAGTATCTCTGCAAAAGTGAATTGGAA 780  
721 Db TTCTGTTCTGATTTGGACACCATGATATAGTTAAGTATCTCTGCAAAAGTGAATTGGAA 780  
781 QY GTTCAACCTCATGTTGTTAATATCTATGAGATACCCCTTACCTGGCATGCTACAAAT 840  
781 Db GTTCAACCTCATGTTGTTAATATCTATGAGATACCCCTTACCTGGCATGCTACAAAT 840  
841 QY GGCATAATTTGAAGTTGCCAAGGAAATCATCCAAATATCAGGAACAGAAAGTCTGACTAAG 900  
841 Db GGCATAATTTGAAGTTGCCAAGGAAATCATCCAAATATCAGGAACAGAAAGTCTGACTAAG 900  
901 QY GAAAAATCTTCACTGTAACAGGTTTTCATAGTGTGTTACCTATGGAAGAGCAATTGAC 960  
901 Db GAAAAATCTTCACTGTAACAGGTTTTCATAGTGTGTTACCTATGGAAGAGCAATTGAC 960  
961 QY CTAGTCAAAATTTCTTCTGATCAGAAATGTCATAAATCATCAACACCAAGGAGGATGGG 1020  
961 Db CTAGTCAAAATTTCTTCTGATCAGAAATGTCATAAATCATCAACACCAAGGAGGATGGG 1020  
1021 QY CACACTGGATTAACCTCTGCTACCGGTTGATCCCGCAGGTCTAGTGGTCAAAATA 1140  
1021 Db CACACTGGATTAACCTCTGCTACCGGTTGATCCCGCAGGTCTAGTGGTCAAAATA 1140  
1081 QY GATAATGGAGCTGATATGATGATCTAGTGGCTTGTGATCCCGCAGGTCTAGTGGTCAAAATA 1140  
1081 Db GATAATGGAGCTGATATGATGATCTAGTGGCTTGTGATCCCGCAGGTCTAGTGGTCAAAATA 1140  
1141 QY GATGAGCAGACATGTTTGTGATGGGCTTATGAAAAAGGCGATGATGCCATTGTCACTC 1200  
1141 Db GATGAGCAGACATGTTTGTGATGGGCTTATGAAAAAGGCGATGATGCCATTGTCACTC 1200  
1201 QY CTGAGACATTAAGAGACCAAGATGATGCTGCTGATCCCGCAGGTCTAGTGGTCAAAATA 1260  
1201 Db CTGAGACATTAAGAGACCAAGATGATGCTGCTGATCCCGCAGGTCTAGTGGTCAAAATA 1260  
1261 QY GGAGATGGCTCCCTATGCTGCTGTTCCATCACCCCTTGGGAGGATTAAGAGATGACAAA 1320  
1261 Db GGAGATGGCTCCCTATGCTGCTGTTCCATCACCCCTTGGGAGGATTAAGAGATGACAAA 1320  
1321 QY GAGAAGGCGATATTTCTTCCCTTAAGAGCTGGATGCTTCAATTTCCATCTTCACTC 1380  
1321 Db GAGAAGGCGATATTTCTTCCCTTAAGAGCTGGATGCTTCAATTTCCATCTTCACTC 1380  
1381 QY TCAGAAATTTGAGTTCCATGAGATTTATGGCTCAGGTTCTTTTGGGAAAGTATATAAGGA 1440  
1381 Db TCAGAAATTTGAGTTCCATGAGATTTATGGCTCAGGTTCTTTTGGGAAAGTATATAAGGA 1440

1441 QY CGATGCGAGAAATAAATAATAGTGGCTATAAAAGCTTATCGAGCCAATACCTACTGCTCAAG 1500  
1441 Db CGATGCGAGAAATAAATAATAGTGGCTATAAAAGCTTATCGAGCCAATACCTACTGCTCAAG 1500  
1501 QY TCAGATGCGATATGCTTTTGGCGAGAGGTGCCATCTCTGCCAGCTCAATCATCCCTGC 1560  
1501 Db TCAGATGCGATATGCTTTTGGCGAGAGGTGCCATCTCTGCCAGCTCAATCATCCCTGC 1560  
1561 QY GTAATTCAGTTTGTGGTGGTCTTGAATGATCCAGCCAGTTTGGCCATTGTCACTCAA 1620  
1561 Db GTAATTCAGTTTGTGGTGGTCTTGAATGATCCAGCCAGTTTGGCCATTGTCACTCAA 1620  
1621 QY TACATATCAGGGGGTCTCTGTTCTCCCTCTCATGAGCAGAGAGGATTTCTTGATTTG 1680  
1621 Db TACATATCAGGGGGTCTCTGTTCTCCCTCTCATGAGCAGAGAGGATTTCTTGATTTG 1680  
1681 QY CAGTCTAAATTAATTAATGTCAGTATGTTGCCAAGGCAATGAGTACCTTTCAACAACCTG 1740  
1681 Db CAGTCTAAATTAATTAATGTCAGTATGTTGCCAAGGCAATGAGTACCTTTCAACAACCTG 1740  
1741 QY ACAGGCCAATTAATCATGCTGACCTTGAAGTGAACAGTCAATATTTCTCTATGAGGATGG 1800  
1741 Db ACAGGCCAATTAATCATGCTGACCTTGAAGTGAACAGTCAATATTTCTCTATGAGGATGG 1800  
1801 QY CATGCTGTGTCGTCAGATTTTGGAGAATCAAGATTTCTACAGTCTCTGATGAAGAACAAC 1860  
1801 Db CATGCTGTGTCGTCAGATTTTGGAGAATCAAGATTTCTACAGTCTCTGATGAAGAACAAC 1860  
1861 QY ATGACAAAAACCTGGGAACTCCGTTGGATGGCTCTGAGGTTTCCAGCAGTGCAT 1920  
1861 Db ATGACAAAAACCTGGGAACTCCGTTGGATGGCTCTGAGGTTTCCAGCAGTGCAT 1920  
1921 QY CGGTACACCATCAAGCAGATGCTTACAGTATGCTCTGCTGTGCTGTGGGAAATTTCTCACT 1980  
1921 Db CGGTACACCATCAAGCAGATGCTTACAGTATGCTCTGCTGTGCTGTGGGAAATTTCTCACT 1980  
1981 QY GGGGAAATTTCCATTTGGCTCATCTCAAGCCAGCGGCTGGCGCAGCAGACATGGCTTACCAC 2040  
1981 Db GGGGAAATTTCCATTTGGCTCATCTCAAGCCAGCGGCTGGCGCAGCAGACATGGCTTACCAC 2040  
2041 QY CACATCAGACCTCCCATTTGGCTATTCCATTTCCAAAGCCCATATCATCTCTGCTGATAGA 2100  
2041 Db CACATCAGACCTCCCATTTGGCTATTCCATTTCCAAAGCCCATATCATCTCTGCTGATAGA 2100  
2101 QY GGTGGAACGCTATGCTCTCAAGGAAGACCCGAATTTCTGAAATTTGTATGAAGTTAGAA 2160  
2101 Db GGTGGAACGCTATGCTCTCAAGGAAGACCCGAATTTCTGAAATTTGTATGAAGTTAGAA 2160  
2161 QY GAGTGTCTCTGCAACATTCAGCTGATGCTCTGCAATCAAGTAAACAGCAGTGGGCTCTCT 2220  
2161 Db GAGTGTCTCTGCAACATTCAGCTGATGCTCTGCAATCAAGTAAACAGCAGTGGGCTCTCT 2220  
2221 QY TCACCTTCTCTTCTCTGATTTGGCTGTAACCGGGAGGACCTGGCGGAGTCAATGTG 2280  
2221 Db TCACCTTCTCTTCTCTGATTTGGCTGTAACCGGGAGGACCTGGCGGAGTCAATGTG 2280  
2281 QY GCAGCTTAAGAGTGGTCTTGAATTTGGAATATGCTCTAAATGCAAGGTCCTATGCTGCT 2340  
2281 Db GCAGCTTAAGAGTGGTCTTGAATTTGGAATATGCTCTAAATGCAAGGTCCTATGCTGCT 2340  
2341 QY TTGTCCCAAGTGGCTGGACAAATATTTCTCTCAAGCTGCTCTTTGCGAGGAGATGAAGA 2400  
2341 Db TTGTCCCAAGTGGCTGGACAAATATTTCTCTCAAGCTGCTCTTTGCGAGGAGATGAAGA 2400  
2401 QY AGTCTTCAATACACACCCATTGCAAAATATGGCTATGATCCGATCCCATGAGCTCAATG 2460  
2401 Db AGTCTTCAATACACACCCATTGCAAAATATGGCTATGATCCGATCCCATGAGCTCAATG 2460  
2461 QY CATTTTCATTTTGGCGAAATATGATAGCAGCTTTGAGGACAGCAGC 2505  
2461 Db CATTTTCATTTTGGCGAAATATGATAGCAGCTTTGAGGACAGCAGC 2505

RESULT 4  
AX056437  
LOCUS AX056437 2508 bp DNA linear PAT 13-JAN-2001  
DEFINITION Sequence 81 from Patent WO0073469.  
ACCESSION AX056437  
VERSION AX056437.1 GI:12229144

KEYWORDS Homo sapiens (human)

SOURCE  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE  
1 Plozman, G.D., Martinez, R., Whyte, D. and Sudersanam, S.

AUTHORS Plozman, G.D., Martinez, R., Whyte, D. and Sudersanam, S.  
TITLE Protein kinases  
JOURNAL Patent: WO 0073469-A 81 07-DEC-2000;  
Sugen, Inc. (US)

FEATURES  
source

Location/Qualifiers  
1..2508  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

ORIGIN

Query Match 100.0%; Score 2505; DB 6; Length 2508;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ATGGGAATTATAATCTAGACCAACCCAACTTGACTGATGAATGGAGAAAAGTC 60
DB 1 ATGGGAATTATAATCTAGACCAACCCAACTTGACTGATGAATGGAGAAAAGTC 60
QY 61 AGTGAATCATATGTTATCACAAATAGAAAGATTAGAGATGACCTGCAGATCAAGGAAAA 120
DB 61 AGTGAATCATATGTTATCACAAATAGAAAGATTAGAGATGACCTGCAGATCAAGGAAAA 120
QY 121 GAATGACAGAACTAAGAAATATATTTGGCTCTGATGAGAGCTTCAGTAAAGTCAATTTA 180
DB 121 GAATGACAGAACTAAGAAATATATTTGGCTCTGATGAGAGCTTCAGTAAAGTCAATTTA 180
QY 181 AATTACCGCACTGAAATGGCTGCTCTACTTCAATTTAGTTGCAATTTGGAGGCAAG 240
DB 181 AATTACCGCACTGAAATGGCTGCTCTACTTCAATTTAGTTGCAATTTGGAGGCAAG 240
QY 241 AAATCACATATTCGAATCTTTATGTTGAAAGGGCTCGGCCATCTCGACTGCAAGAAAT 300
DB 241 AAATCACATATTCGAATCTTTATGTTGAAAGGGCTCGGCCATCTCGACTGCAAGAAAT 300
QY 301 GGATTTACAGCTTTGCATTTAGCAGTTTACAGGATATGCAAGATTTGATCTCTCTG 360
DB 301 GGATTTACAGCTTTGCATTTAGCAGTTTACAGGATATGCAAGATTTGATCTCTCTG 360
QY 361 CTTACAGTGGAGCTGATATACAGAGCTTGGATACGGTGGCTCACTGCCCTCCATATT 420
DB 361 CTTACAGTGGAGCTGATATACAGAGCTTGGATACGGTGGCTCACTGCCCTCCATATT 420
QY 421 GCTACATAGTGGCCACCTAGAGGCTGCTGATGCTGTTGCAATGAGAGCTAATGTC 480
DB 421 GCTACATAGTGGCCACCTAGAGGCTGCTGATGCTGTTGCAATGAGAGCTAATGTC 480
QY 481 AATATTCAAGATGAGTTTTTTTCACTCCATTCATATTTGCAAGCTGCTGATGATGAA 540
DB 481 AATATTCAAGATGAGTTTTTTTCACTCCATTCATATTTGCAAGCTGCTGATGATGAA 540
QY 541 CAGTAACTCGCTCTCTTTTGAATTTGGTGTGATGTAATTTGAGTGGTGAAGTTGGA 600
DB 541 CAGTAACTCGCTCTCTTTTGAATTTGGTGTGATGTAATTTGAGTGGTGAAGTTGGA 600
QY 601 GATAGACCTTCCACCTAGCATCTGCAAAAGGATCTTGAATTTGCAAACTCTTGATG 660
DB 601 GATAGACCTTCCACCTAGCATCTGCAAAAGGATCTTGAATTTGCAAACTCTTGATG 660
QY 661 GAAGAGGCGAGAAAGCAGATGTAATGCTCAAGATAATGAAGACCATGTCCCATCCAT 720
DB 661 GAAGAGGCGAGAAAGCAGATGTAATGCTCAAGATAATGAAGACCATGTCCCATCCAT 720
```

```
QY 1801 CATGCTGTGGTGGCAGATTTTCGAGAACTCAAGATTTCTACAGTCTCTCGATGAAGACAAC 1860
Db 1801 CATGCTGTGGTGGCAGATTTTCGAGAACTCAAGATTTCTACAGTCTCTCGATGAAGACAAC 1860
QY 1861 ATGACAAAACAACTGGGAACTCCGTTGGATGGCTTCCTGAGGTTCACCGAGTGCAC 1920
Db 1861 ATGACAAAACAACTGGGAACTCCGTTGGATGGCTTCCTGAGGTTCACCGAGTGCAC 1920
QY 1921 CGGTACACCATCAAGACAGATCTCTTCAGTATGCTCTGCTGCTGCTGGAAATTCCTCACT 1980
Db 1921 CGGTACACCATCAAGACAGATCTCTTCAGTATGCTCTGCTGCTGCTGGAAATTCCTCACT 1980
QY 1981 GCGGAAATTCATTCCTCTCATCTCAAGCAGCGGCTGGCGCAGACAGACATGGCTTACCA 2040
Db 1981 GCGGAAATTCATTCCTCTCATCTCAAGCAGCGGCTGGCGCAGACAGACATGGCTTACCA 2040
QY 2041 CACATCAGACCTCCATTCGCTATTCCTTCCATTCGAGCCCATATCATCTCTGCTGATACGA 2100
Db 2041 CACATCAGACCTCCATTCGCTATTCCTTCCATTCGAGCCCATATCATCTCTGCTGATACGA 2100
QY 2101 GGGTGAACGCGATGCTCAAGGAGACACCGAATTTCTGAAGTTGTCATGAAGTTAGAA 2160
Db 2101 GGGTGAACGCGATGCTCAAGGAGACACCGAATTTCTGAAGTTGTCATGAAGTTAGAA 2160
QY 2161 GAGTGTCTCTGCAACATTCAGCTGATGCTCTCTGATCAAGTAACAGCAGTGGGTCTCTC 2220
Db 2161 GAGTGTCTCTGCAACATTCAGCTGATGCTCTCTGATCAAGTAACAGCAGTGGGTCTCTC 2220
QY 2221 TCACCTCTCTCTCTCTCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2280
Db 2221 TCACCTCTCTCTCTCTCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2280
QY 2281 GCAGCATTAAGAACTGCTTTCGAAATGGAATATGCTTAAATGCAAGTCTCTGCTGCT 2340
Db 2281 GCAGCATTAAGAACTGCTTTCGAAATGGAATATGCTTAAATGCAAGTCTCTGCTGCT 2340
QY 2341 TTGTCGCAAGTCTGCAAAATTCCTCTCAAGTCTGCTTTCGAGAGATGGAAGA 2400
Db 2341 TTGTCGCAAGTCTGCAAAATTCCTCTCAAGTCTGCTTTCGAGAGATGGAAGA 2400
QY 2401 AGTCTTCAATACACACCATGACAAATATGCTATGCTATGCTATGCTATGCTATGCTATGCT 2460
Db 2401 AGTCTTCAATACACACCATGACAAATATGCTATGCTATGCTATGCTATGCTATGCTATGCT 2460
QY 2461 CATTTCATTCCTGCGAAATAGTACAGCTTTGAGGACAGCAGC 2505
Db 2461 CATTTCATTCCTGCGAAATAGTACAGCTTTGAGGACAGCAGC 2505

RESULT 5
AY303691 2980 bp mRNA linear PRI 23-JUN-2003
LOCUS Homo sapiens cardiac ankyrin repeat kinase mRNA, complete cds.
DEFINITION
ACCESSION AV303691
VERSION AV303691.1 GI:32165605
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2980)
AUTHORS Jeyaseelan,R.
TITLE Cardiac Ankyrin Repeat Kinase (CARK)
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2980)
AUTHORS Jeyaseelan,R.
TITLE Direct Submission
JOURNAL Submitted (20-MAY-2003) Cardiovascular Biology, Millennium
Pharmaceuticals Inc., 75 Sidney Street, Cambridge, MA 02139, USA
FEATURES
source Location/Qualifiers
1..2980
/organism="Homo sapiens"
```

```
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="1"
/map="1p22.3-p13.3"
31..2538
/note="CARK; protein kinase"
/codon_start=1
/product="cardiac ankyrin repeat kinase"
/protein_id="AAP72030.1"
/db_xref="GI:32165606"
/translation="MGNYKSRPTCTDDEWKKVSVYITIERLEDDLOIKEXELTE
LENIFGSEARSKVNLNRYTENGSLHLCOICGKSKSHRTLMKGLRPSRLTRNGP
TALHAYTKDNALIELSLHSAGDIQQVGYGGLTALHIAITAGLEADVLLQHGAVU
NIQDAVFTPLHIAAYGHEQVTRLLKFGADVNVSGEVRPDLHLASAKGFLNIAKL
LMEEGSKADVNAQDNEDHVLPHFSRFGHDIKYLQSLQDELVFVHVNIYDTPDLH
ACYNGKFEVAKELIIQISGTELSKXENIFSETAFHSACTYKSIDLVKFLDQVININ
HOGDRGHTGLHSACVGHIRLVQFLDNGADMTVACDPSRSGEKDEQTCCLMAYEK
GHDALVTGLKVKRQDELPCNEYSGPGDGSVYVSPGLKIKSMYKEXADILLRRA
GLPSFHILQLEISEIPEHEIIGSGSGKVKYKGRNKIYALKYRANTVCSKSDVDMFCR
EYVILCQNLNCPVIOIQAFLNLSNILLYEDGHVAVDPEGRFLLDLOLQSKLII
AYDVAKGMEYDHLNLTQPIIHRDLNSHLLYEDGHVAVDPEGRFLLDLOLQSKLII
PGLNLMAPVPTQCTRYTIIKADVFSYALCWEILTEIIPFAHLKPAADAAWYHHI
RPIGVSYIPKIPISLLIRGMNACPEGRPEFSEVVMKLEECLENTLEMSPASNSGSL
SPSSSDCLVNRGGPGRSHVAALRSRPELEYALNARSVAALSQSGQVSSQGLSLEBM
KRSLOYTPIDKYGVYSDPMSSMHFHSNRSNRPEDSS"
```

ORIGIN

QY	1	ATGGGAAATATTAATCTAGACCAACCCAACTTGTACTGATGAATGAAGAAAAAGTC	60
Db	31	ATGGGAAATATTAATCTAGACCAACCCAACTTGTACTGATGAATGAAGAAAAAGTC	90
QY	61	AGTGAATCATATGTTATCAATAGAAAGATTAGAAGATGACCTGCAGATCAAGAAAAA	120
Db	91	AGTGAATCATATGTTATCAATAGAAAGATTAGAAGATGACCTGCAGATCAAGAAAAA	150
QY	121	GAATCAGACAACTAAGGATATATTTGGCTCTGATGAAGCTTCAGTAAAGTCAATTTA	180
Db	151	GAATCAGACAACTAAGGATATATTTGGCTCTGATGAAGCTTCAGTAAAGTCAATTTA	210
QY	181	AATTACCGCACTGAAATGGGCTGCTCTACTTCAATTTATGTTGATTTGGAGGCAAG	240
Db	211	AATTACCGCACTGAAATGGGCTGCTCTACTTCAATTTATGTTGATTTGGAGGCAAG	270
QY	241	AAATCACAATTCGAACCTTTATGTTGAAAGGGCTCCGCCATCTCGACTGACAGAAAT	300
Db	271	AAATCACAATTCGAACCTTTATGTTGAAAGGGCTCCGCCATCTCGACTGACAGAAAT	330
QY	301	GGATTTACAGCTTCGATTTAGCAGTTTACAGGATAATGCAGATTTGATCACTTCTCTG	360
Db	331	GGATTTACAGCTTCGATTTAGCAGTTTACAGGATAATGCAGATTTGATCACTTCTCTG	390
QY	361	CTTCACAGTGGAGCTGATATACAGCAGGTTGGATACGGTGGCTCACTGCCCTCCATAT	420
Db	391	CTTCACAGTGGAGCTGATATACAGCAGGTTGGATACGGTGGCTCACTGCCCTCCATAT	450
QY	421	GCTACAAATAGTGGCCACCTAGAGCTGCTGATGCTGTTGTCACATGGAGCTAATGTC	480
Db	451	GCTACAAATAGTGGCCACCTAGAGCTGCTGATGCTGTTGTCACATGGAGCTAATGTC	510
QY	481	AAATATTCAAGATGCGAGTTTTTTTCTACTCCATTCGATATTTGCAGCGTACTATGGACATGAA	540
Db	511	AAATATTCAAGATGCGAGTTTTTTTCTACTCCATTCGATATTTGCAGCGTACTATGGACATGAA	570
QY	541	CAGTAACTCGCTCTTTTGTGAAATTTGGTGTGATGTAATGTAAGTGGTGAAGTTGGA	600
Db	571	CAGTAACTCGCTCTTTTGTGAAATTTGGTGTGATGTAATGTAAGTGGTGAAGTTGGA	630
QY	601	GATAGACCCCTCCACCTAGCATCTGCAGAAAGGATTTCTGAATATTGCAAAACTCTTGATG	660

Df	631	GATAGACCCCTCCACCTAGCATCTGCAAAAGGATTCTTGAATATTCGAAACCTCTTGATG	690
QY	661	GAAGAAGCAGCAAGCAGATGTGAATGCTCAAGATAATGAAGACCATGTCCTCCAT	720
Df	691	GAAGAAGCAGCAAGCAGATGTGAATGCTCAAGATAATGAAGACCATGTCCTCCAT	750
QY	721	TTCTCTCTCGAATTTGGACACCATGATATAGTTAAGTATCTGCTGCAAAAGTATTTGGAA	780
Df	751	TTCTCTCTCGAATTTGGACACCATGATATAGTTAAGTATCTGCTGCAAAAGTATTTGGAA	810
QY	781	GTTCACCTCATGTTGTTTAATATCTATGGAGATACCCCTTTACACCTGGCATCTACAAT	840
Df	811	GTTCACCTCATGTTGTTTAATATCTATGGAGATACCCCTTTACACCTGGCATCTACAAT	870
QY	841	GGCAATTTGAAGTGGCAAGGAATCATCCAAATATCAGAACAGAAAGTCTGACTAG	900
Df	871	GGCAATTTGAAGTGGCAAGGAATCATCCAAATATCAGAACAGAAAGTCTGACTAG	930
QY	901	CAAAAACATCTTCAGTGAACACAGCTTTTTCATAGTCTTTGTACCTATGCGAAGACATTGAC	960
Df	931	CAAAAACATCTTCAGTGAACACAGCTTTTTCATAGTCTTTGTACCTATGCGAAGACATTGAC	990
QY	961	CTAGTCAAAATTTCTTCTTGATCAGAAATGTCATAAATCAACCAAGCAAGGATGG	1020
Df	991	CTAGTCAAAATTTCTTCTTGATCAGAAATGTCATAAATCAACCAAGCAAGGATGG	1050
QY	1021	CACACTGATTTACACTCTGCTTGCTACACGGTCAACATTCGCTGGTTCAGTTCTTACTG	1080
Df	1051	CACACTGATTTACACTCTGCTTGCTACACGGTCAACATTCGCTGGTTCAGTTCTTACTG	1110
QY	1081	GATATGAGCTGATATGAATCTAGTGGCTTGATCCAGAGAGTCTAGTGGTGAATA	1140
Df	1111	GATATGAGCTGATATGAATCTAGTGGCTTGATCCAGAGAGTCTAGTGGTGAATA	1170
QY	1141	GATAGCAGACATGTTGATGGGCTTATGAAGAGGCGATGCGATGTCACATC	1200
Df	1171	GATAGCAGACATGTTGATGGGCTTATGAAGAGGCGATGCGATGTCACATC	1230
QY	1201	CTGAGCATTATAAGACACCAAGATGAATGCTGTAATGAATATTTCTCAGCTGGA	1260
Df	1231	CTGAGCATTATAAGACACCAAGATGAATGCTGTAATGAATATTTCTCAGCTGGA	1290
QY	1261	GGAGTGGCTCTATGCTGCTGCTCCTACACCTTGGGAGAGTTAAAGCATGACAAA	1320
Df	1291	GGAGTGGCTCTATGCTGCTGCTCCTACACCTTGGGAGAGTTAAAGCATGACAAA	1350
QY	1321	GAGAGGCGATATTTCTCTCTAAGAGCTGGATTCCTTCAATTTCCATCTTCAGTCT	1380
Df	1351	GAGAGGCGATATTTCTCTCTAAGAGCTGGATTCCTTCAATTTCCATCTTCAGTCT	1410
QY	1381	TCAGAAATGAGTTCATGAGATTAATGGCTCAGGTTCTTTGGGAAAGTATATAAGGA	1440
Df	1411	TCAGAAATGAGTTCATGAGATTAATGGCTCAGGTTCTTTGGGAAAGTATATAAGGA	1470
QY	1441	CGATGCAAAATAAATAGTGGCTATATAAAGCTTATCGAGCCAAATACCTACTGCTCAAG	1500
Df	1471	CGATGCAAAATAAATAGTGGCTATATAAAGCTTATCGAGCCAAATACCTACTGCTCAAG	1530
QY	1501	TCAGATGGAATATGTTTTCGAGAGGCTGCTCAATCTCTGCGAGTCAATCATCCCTGC	1560
Df	1531	TCAGATGGAATATGTTTTCGAGAGGCTGCTCAATCTCTGCGAGTCAATCATCCCTGC	1590
QY	1561	GTAATTCAGTTTGGTGGTCTGCTGAATGATCCAGCCAGTTTCCATTTGCTCACTCA	1620
Df	1591	GTAATTCAGTTTGGTGGTCTGCTGAATGATCCAGCCAGTTTCCATTTGCTCACTCA	1650
QY	1621	TACATATCAGGGGTTCTGTTCTCCCTCTTCATGAGCAAGAGGATTTCTGATTG	1680
Df	1651	TACATATCAGGGGTTCTGTTCTCCCTCTTCATGAGCAAGAGGATTTCTGATTG	1710
QY	1681	CAGTCTAAATTAATTTGAGTAGATGTTGCCAAAGGATGAGTACCTTCACAACTG	1740
Df	1711	CAGTCTAAATTAATTTGAGTAGATGTTGCCAAAGGATGAGTACCTTCACAACTG	1770
QY	1741	ACACAGCAATTTATATCATGCTGACTTGAACAGTCACAATATTTCTTCTCTAGGATGG	1800
Df	1771	ACACAGCAATTTATATCATGCTGACTTGAACAGTCACAATATTTCTTCTCTAGGATGG	1830
QY	1801	CATCTCTGCTGGCAGATTTTGGAGAAATCAAGATTTCTACAGTCTCTGGATGAACAAC	1860
Df	1831	CATCTCTGCTGGCAGATTTTGGAGAAATCAAGATTTCTACAGTCTCTGGATGAACAAC	1890
QY	1861	ATGACAAAACAACTGGAAACCTCCCTTGGATGGCTCTCTAGAGTGTTCAGCAGTGCAT	1920
Df	1891	ATGACAAAACAACTGGAAACCTCCCTTGGATGGCTCTCTAGAGTGTTCAGCAGTGCAT	1950
QY	1921	CGGTACACCATCAAGCAGATGCTTCAGCTATGCTCTGCTGTCTGGAAATTTCTCACT	1980
Df	1951	CGGTACACCATCAAGCAGATGCTTCAGCTATGCTCTGCTGTCTGGAAATTTCTCACT	2010
QY	1981	GGGAAATTCATTCGCTCATCTCAAGCAGCGGCTGGGACAGACATGGCTTACCAAC	2040
Df	2011	GGGAAATTCATTCGCTCATCTCAAGCAGCGGCTGGGACAGACATGGCTTACCAAC	2070
QY	2041	CACATCAGACCTCCCATTCGCTATTCATTCCTCCAGCCCATATCATCTCTGCTGATCGA	2100
Df	2071	CACATCAGACCTCCCATTCGCTATTCATTCCTCCAGCCCATATCATCTCTGCTGATCGA	2130
QY	2101	GGGTGGAAACGATGCTCTGAAGGAAGACCCGAAATTTCTGAAGTTGTCTGAAGTTAGAA	2160
Df	2131	GGGTGGAAACGATGCTCTGAAGGAAGACCCGAAATTTCTGAAGTTGTCTGAAGTTAGAA	2190
QY	2161	GAGTGTCTGCAACATTCAGCTGATGCTCTCTGATCAAGTAAACAGAGTGGTCTCTC	2220
Df	2191	GAGTGTCTGCAACATTCAGCTGATGCTCTCTGATCAAGTAAACAGAGTGGTCTCTC	2250
QY	2221	TCACCTTCTTCTTCTCTGATTCCTGTTGAACCGGGAGGACCTGGCGGAGTCAATG	2280
Df	2251	TCACCTTCTTCTTCTGATTCCTGTTGAACCGGGAGGACCTGGCGGAGTCAATG	2310
QY	2281	GCAGCATTAAGAAAGTCTTGAATTTGAATATGCTCTAAATGGAAGTCTCTATGCTCT	2340
Df	2311	GCAGCATTAAGAAAGTCTTGAATTTGAATATGCTCTAAATGGAAGTCTCTATGCTCT	2370
QY	2341	TTGTCCCAAGTCTGCAATATTCCTCTCAAGTCTGCTTTTGGAGAGATGAAGA	2400
Df	2371	TTGTCCCAAGTCTGCAATATTCCTCTCAAGTCTGCTTTTGGAGAGATGAAGA	2430
QY	2401	AGTCTTCAATACACACCATTCGCAATATGCTATGATTCGATCCCATGAGTCAATG	2460
Df	2431	AGTCTTCAATACACACCATTCGCAATATGCTATGATTCGATCCCATGAGTCAATG	2490
QY	2461	CATTTTCATCTGCGCAATATGACAGCTTTGAGGACAGCAGC	2505
Df	2491	CATTTTCATCTGCGCAATATGACAGCTTTGAGGACAGCAGC	2535
RESULT 6			
LOCUS	AF116826	3001 bp	mRNA linear PRI 11-MAY-1999
DEFINITION	Homo sapiens clone HH498 putative protein-tyrosine kinase mRNA,		
	complete cds.		
ACCESSION	AF116826		
VERSION	AF116826.1	GI:4768828	
KEYWORDS	FLI CDNA.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE			
AUTHORS	Wei, Y.J., Ding, J.F., Xiong, H., Zhou, Y., Hui, R.T. and Liew, C.C.		
TITLE	Submitted (25-DEC-1998) Molecular Medicine Center, Cardiovascular		
JOURNAL	Institute and Pu Wai Hospital, 167 Baileishi Road, West District,		
	Beijing 100037, China		
FEATURES	Location/Qualifiers		

## source

```
1. 3001
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="HH498"
/tissue_type="heart"
/dev_stage="adult"
52. 2559
/codon_start=1
/product="putative protein-tyrosine kinase"
/protein_id="AAD29632.1"
/db_xref="GI:4768829"
```

## CDS

```
/translation="MGNYKSRPTCTDEWKKKVSESVITIERLEDLQIKEKELTE
LNIFGSDAFKSNLVNRTENGLSLHLLCCICGKKSHIRTLMLKLRPSRLTRNGF
TALHVAVKDNLAEILTSLLHSGADIQOYGVGGTALHIAITAGHLEADVLLQHGAV
NCDVAFPTPLHIAAYVGHQVTRLLKFGADVNVSGEVDRLPLHLSAKGFLNIAKL
LMEBSKADVNQDNEDHVPJLHFCRFGHDIIVKYLQSLDLEOVPHVNIYIGDTPHL
ACTNKGFEVAKETIQISGESLITKENIFSETAFHSACTYKGSILDLVKFLDDVININ
HGRDGHGTLSHACHGHIRLVQFLDNGADNLVADPSSSGKEBKEQICLWAYEK
GHDAIVTLKHVRYQDELPCNYSQPGDGSVSVFPLGKIKIMTKEXKADILLRFA
GLPSHFHLQSLSEIEFHEIGSGFGKVKGRNKAIVAKRYRANTYCSKSDVDFCR
EVSILCOLNHPCVIOFVGACLNDSPOFAIVTQYISGSLFSLHLEQRILDLQSKLI
AVDVKGMRYLNLQPIIHRDLNHNLLYEDGHAVVDFEGRFQLSDENMTKQ
PGLNRMAEVTQCTRYTIKADPVSVALCLWEILTGEIPEAHLKPAARAADMAVHI
RPPIGYSIPKISSILLIRGNACPEGRFEFSEVNMKLECLNIELMSPASSNSSL
SPSSSDCLVNGGGRSVAALRFRFEYALNARSYAALSQSAQISSQGLSLEEM
KRSLOVTPIDKTYGVSDPMSMHFHSRNSSEFSDS"
```

polya\_signal

2978. 2983

## ORIGIN

```
Query Match 100.0%; Score 2505; DB 9; Length 3001;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGAAATTAATAACTAGACCAACCCAAACTGTACTGATGAATGGAAGAAAAGTC 60
DB 52 ATGGGAAATTAATAACTAGACCAACCCAAACTGTACTGATGAATGGAAGAAAAGTC 111
QY 61 AGTGAATCATATGTTATCAATAGAAAAGATTAGAAGATGACCTGCAGATCAAGAAAA 120
DB 112 AGTGAATCATATGTTATCAATAGAAAAGATTAGAAGATGACCTGCAGATCAAGAAAA 171
QY 121 GAATGACAGAACTAGGAATATATTTGGCTCTGATGAAGCCCTTCAGTAAAGTCAATTTA 180
DB 172 GAATGACAGAACTAGGAATATATTTGGCTCTGATGAAGCCCTTCAGTAAAGTCAATTTA 231
QY 181 AATTACCGCACTGAAATGGGCTGTCTCTACTTCATTTATGTTGCAATTTGTGGAGGCAAG 240
DB 232 AATTACCGCACTGAAATGGGCTGTCTCTACTTCATTTATGTTGCAATTTGTGGAGGCAAG 291
QY 241 AATACATATTCGAATCTTATTTGAAAGGCTCCGCCATCTCGACTGACAAAGAAAT 300
DB 292 AATACATATTCGAATCTTATTTGAAAGGCTCCGCCATCTCGACTGACAAAGAAAT 351
QY 301 GGATTTACAGCTTGCAATTTAGCAGTTTACAAGATAATGCAAGATTGATCACTTCTCTG 360
DB 352 GGATTTACAGCTTGCAATTTAGCAGTTTACAAGATAATGCAAGATTGATCACTTCTCTG 411
QY 361 CTTACAGTGGAGCTGATATACAGCAGGTGGATAGCGTGGCTCTACTGCCCTCATATT 420
DB 412 CTTACAGTGGAGCTGATATACAGCAGGTGGATAGCGTGGCTCTACTGCCCTCATATT 471
QY 421 GCTCAATAGCTGCCACCTAGAGGCTGCTGATGCTGTGCTTGCAACATGGAGCTAATGTC 480
DB 472 GCTCAATAGCTGCCACCTAGAGGCTGCTGATGCTGTGCTTGCAACATGGAGCTAATGTC 531
QY 481 AATATTCAAGATGCAAGTTTTTTTCACTCCATTCATATTCGACGGTACATGACATGAA 540
DB 532 AATATTCAAGATGCAAGTTTTTTTCACTCCATTCATATTCGACGGTACATGACATGAA 591
QY 541 CAGTAACTCCCTCTTTTGAATTTGGTCTGATGCTAAATGTAAGTGTGAGTTGGA 600
DB 592 CAGTAACTCCCTCTTTTGAATTTGGTCTGATGCTAAATGTAAGTGTGAGTTGGA 651
```

QY 1681 CAGTCTAAATTAATTTAGTAGATGTTGCCAAGGATGAGTACCTTTCACAACTG 1740  
Db 1732 CAGTCTAAATTAATTTAGTAGATGTTGCCAAGGATGAGTACCTTTCACAACTG 1791  
QY 1741 ACACAGCCAAATTAATTCAGTGTGAACAGTACCAATATTTCTCTATGAGGATGG 1800  
Db 1792 ACACAGCCAAATTAATTCAGTGTGAACAGTACCAATATTTCTCTATGAGGATGG 1851  
QY 1801 CATGCTGTGGTGGAGATTTGGAGATCAAGATTTCTACAGTCTCTGGATGAGCAAC 1860  
Db 1852 CATGCTGTGGTGGAGATTTGGAGATCAAGATTTCTACAGTCTCTGGATGAGCAAC 1911  
QY 1861 ATGACAAAACCACTGGGAACCTCGTTGGATGGCTCTGAGGTGTTTCACGCCAGTGCAC 1920  
Db 1912 ATGACAAAACCACTGGGAACCTCGTTGGATGGCTCTGAGGTGTTTCACGCCAGTGCAC 1971  
QY 1921 CGGTACACCAATCAAGAGATGTTCTGAGTATGCTCTGTCTGTGGGAAATTCAC 1980  
Db 1972 CGGTACACCAATCAAGAGATGTTCTGAGTATGCTCTGTCTGTGGGAAATTCAC 2031  
QY 1981 GCGGAAATTCATTCGCTCATCTCAAGCCAGCGCTCGGAGCAGACATGGCTTACCAC 2040  
Db 2032 GCGGAAATTCATTCGCTCATCTCAAGCCAGCGCTCGGAGCAGACATGGCTTACCAC 2091  
QY 2041 CACATCAGACCTCCCAATGGCTATTCATTCCTCCAAAGCCCATATCATCTCTGTGATACGA 2100  
Db 2092 CACATCAGACCTCCCAATGGCTATTCATTCCTCCAAAGCCCATATCATCTCTGTGATACGA 2151  
QY 2101 GGGTGGAGCGATGCTCTGAGGAGACCCGAAATTTCTGAAGTGTCTCATGAGTTAGAA 2160  
Db 2152 GGGTGGAGCGATGCTCTGAGGAGACCCGAAATTTCTGAAGTGTCTCATGAGTTAGAA 2211  
QY 2161 GAGTGTCTCTCAACATTTGAGTGTCTCTGATCAATCAAGTAAAGCAGTGGTCTCTC 2220  
Db 2212 GAGTGTCTCTCAACATTTGAGTGTCTCTGATCAATCAAGTAAAGCAGTGGTCTCTC 2271  
QY 2221 TCACCTCTCTCTCTCTGATGCTGTGTAACCGGGAGACCTGGCCGGAGTCAATG 2280  
Db 2272 TCACCTCTCTCTCTCTGATGCTGTGTAACCGGGAGACCTGGCCGGAGTCAATG 2331  
QY 2281 GCAGCATTAAGAAGTCTGTTTGGAAATGGAATATGCTCTAAATCAAGGTCTTATGCTCT 2340  
Db 2332 GCAGCATTAAGAAGTCTGTTTGGAAATGGAATATGCTCTAAATCAAGGTCTTATGCTCT 2391  
QY 2341 TTGTCCCAAGTGTGACAAATATTCCTCTCAAGGTCTGTCTTTGGAGAGATGAAAGA 2400  
Db 2392 TTGTCCCAAGTGTGACAAATATTCCTCTCAAGGTCTGTCTTTGGAGAGATGAAAGA 2451  
QY 2401 AGTCTTCAATACACACCACTTGACAAATATGCTATGCTATCCGATCCCATGAGCTCAATG 2460  
Db 2452 AGTCTTCAATACACACCACTTGACAAATATGCTATGCTATCCGATCCCATGAGCTCAATG 2511  
QY 2461 CATTTTCATTTCCGGAATAGTAGAGCTTTTGGAGCAGCAGC 2505  
Db 2512 CATTTTCATTTCCGGAATAGTAGAGCTTTTGGAGCAGCAGC 2556

RESULT 7  
BD269559 3025 bp DNA linear PAR 17-JUL-2003  
LOCUS BD269559  
DEFINITION Novel CARP protein and nucleic acid molecules and uses therefor.  
ACCESSION BD269559  
VERSION BD269559.1 GI:33079327  
KEYWORDS JP 2002535962-A/1.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 3025)  
AUTHORS Raju, J.  
TITLE Novel CARP protein and nucleic acid molecules and uses therefor  
JOURNAL Patent: JP 2002535962-A 1 29-OCT-2002;  
MILLENNIUM PHARMACEUTICALS INC

COMMENT OS Homo sapiens (human)  
PN JP 2002535962-A/1  
PD 29-OCT-2002  
PF 10-DEC-1999 JP 2000586772  
PR 11-DEC-1998 US 60/111938,14-APR-1999 US 09/291839 P1  
JEYASEELAN RAJU  
PC C12N15/09,C07K16/40,C12N1/15,C12N1/19,C12N1/21,C12N5/10,C12N9/  
10,C1201/02,  
PC C12Q1/68,G01N33/15,G01N33/50,G01N33/53,G01N33/68//A61K38/45,  
PC A61K39/395,  
PC A61K45/00,A61P9/04,A61P9/10,A61P35/00,A61P43/00,C12N15/00, PC  
C12N5/00,  
PC A61K37/52  
CC Novel CARP protein and nucleic acid molecules and uses CC  
FH Key Location/Qualifiers  
FT CDS (48)..(2552).  
FEATURES  
source  
ORIGIN  
Query Match 100.0%; Score 2505; DB 6; Length 3025;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGGGAATTAATAATCTAGACCAACCCAACTTGTACTGATGAATGAAGAAAATGTC 60  
Db 48 ATGGGAATTAATAATCTAGACCAACCCAACTTGTACTGATGAATGAAGAAAATGTC 107  
QY 61 AGTGAATCATATGTTATCAATAGAAAGATTAGAAAGTACCTCGACATCAAGGAAAA 120  
Db 108 AGTGAATCATATGTTATCAATAGAAAGATTAGAAAGTACCTCGACATCAAGGAAAA 167  
QY 121 GAATGACAGAACTAAGGAATATATTTGGCTCTGATGAAGCCCTCAGTAAAGTCAATTA 180  
Db 168 GAATGACAGAACTAAGGAATATATTTGGCTCTGATGAAGCCCTCAGTAAAGTCAATTA 227  
QY 181 AATTACCGCACTGAAATGGGTCTCTCTACTTCAATTTATGTTGATTTGTTGGAGCAAG 240  
Db 228 AATTACCGCACTGAAATGGGTCTCTCTACTTCAATTTATGTTGATTTGTTGGAGCAAG 287  
QY 241 AATATCAATATTCGAATCTTTATGTTGAAAGGGCTCCGCCATCTCGACTGACAGAAAT 300  
Db 288 AATATCAATATTCGAATCTTTATGTTGAAAGGGCTCCGCCATCTCGACTGACAGAAAT 347  
QY 301 GGATTTACAGCCCTTGCATTTAGCAGTTTACAGGATATGAGGATTCATCAGTCTCTCTG 360  
Db 348 GGATTTACAGCCCTTGCATTTAGCAGTTTACAGGATATGAGGATTCATCAGTCTCTCTG 407  
QY 361 CTTTACAGTGGAGCTGATATACAGCAGGTTGGATACGGTGGCTCACTGCCCTCCATATT 420  
Db 408 CTTTACAGTGGAGCTGATATACAGCAGGTTGGATACGGTGGCTCACTGCCCTCCATATT 467  
QY 421 GCTTACAAATAGTGGCCACCTAGAGCTGCTGATGCTGTTGCAACATGGAGCTAATGTC 480  
Db 468 GCTTACAAATAGTGGCCACCTAGAGCTGCTGATGCTGTTGCAACATGGAGCTAATGTC 527  
QY 481 AATATTCAAGATGCACTTTTTTTTCACTGATATGAGGATTCATGAGGATGATGAGCAATGAA 540  
Db 528 AATATTCAAGATGCACTTTTTTTTCACTGATATGAGGATTCATGAGGATGATGAGCAATGAA 587  
QY 541 CAGGTAACCTGCTCTTTTGAATTTGGTCTGATGTTAAATGTTAAGTGGTGAAGTTGGA 600  
Db 588 CAGGTAACCTGCTCTTTTGAATTTGGTCTGATGTTAAATGTTAAGTGGTGAAGTTGGA 647  
QY 601 GATAGACCCCTCCACTAGCATCTGCAAAAGGATTTCTTGAATATTGCAAACTCTCTGATG 660  
Db 648 GATAGACCCCTCCACTAGCATCTGCAAAAGGATTTCTTGAATATTGCAAACTCTCTGATG 707  
QY 661 GAAGAGGCGCAAGAGCAGATGTGAATGCTCAAGATAATGAAGACCATGTCTCCACTCCAT 720

708	DB	GAAGAGGACGACAAAGCAGATGTGAATGCTCAAGATAATGAAGACCAATGTGCCACCTCCAT	767
721	QY	TTCTGTCTCGATTTGGACACCATGATATAGTTTAAGTATCTCTGCGAAAGTGAATTTGGAA	780
768	DB	TTCTGTCTCGATTTGGACACCAAGATATAGTTTAAGTATCTCTGCGAAAGTGAATTTGGAA	827
781	QY	GTTCAAGCTCATGTGTGTTAATATCTATGAGATACCCCTTTACACCTGGCATGCTACAAT	840
828	DB	GTTCAAGCTCATGTGTGTTAATATCTATGAGATACCCCTTTACACCTGGCATGCTACAAT	887
841	QY	GGCAAAATTGAAGTTGCCAAGGAAATCATCCAAATATCAGGAAACAGAAAGCTTCGACTAAG	900
888	DB	GGCAAAATTGAAGTTGCCAAGGAAATCATCCAAATATCAGGAAACAGAAAGCTTCGACTAAG	947
901	QY	CAAAACATCTTCAGTGAAGAACAGCTTTTCATAGTGGTTGTCCTATGCGACAGACATTCGAC	960
948	DB	GAAACATCTTCAGTGAAGAACAGCTTTTCATAGTGGTTGTCCTATGCGACAGACATTCGAC	1007
961	QY	CTAGTCAAAATTTCTTCTGTGATCAGAATGTCATAAACATCAACCAAGGATGGG	1020
1008	DB	CTAGTCAAAATTTCTTCTGTGATCAGAATGTCATAAACATCAACCAAGGATGGG	1067
1021	QY	CACACTGGATTACACTCTGCTTCTCTACACGGTTCACATTCGCTGGTTTCAGTCTTACTG	1080
1068	DB	CACACTGGATTACACTCTGCTTCTCTACACGGTTCACATTCGCTGGTTTCAGTCTTACTG	1127
1081	QY	GATTAATGGAGCTGATATGAATCTAGTGGCTTGTGATTCGACGAGCTTAGTGTGAAGAA	1140
1128	DB	GATTAATGGAGCTGATATGAATCTAGTGGCTTGTGATTCGACGAGCTTAGTGTGAAGAA	1187
1141	QY	GATGAGCAGACATGTTTGATGTGGGCTTATGAAAAGGGCATGATGCCATGTGCACATC	1200
1188	DB	GATGAGCAGACATGTTTGATGTGGGCTTATGAAAAGGGCATGATGCCATGTGCACATC	1247
1201	QY	CTGAAGCATTTAAGAGACACCAAGATGAATTCGCTGTATGTAATTTCTCAGCTTGGGA	1260
1248	DB	CTGAAGCATTTAAGAGACACCAAGATGAATTCGCTGTATGTAATTTCTCAGCTTGGGA	1307
1261	QY	GGAGATGGCTCCTATGTCTGTTCATCACCTTGGGGAAGATTAAAGCATGACAAAA	1320
1308	DB	GGAGATGGCTCCTATGTCTGTTCATCACCTTGGGGAAGATTAAAGCATGACAAAA	1367
1321	QY	GAGAAGGCAGATATTCCTCCTCTAAGAGCTGGATTCCTTCAATTTCCATTTCACTC	1380
1368	DB	GAGAAGGCAGATATTCCTCCTCTAAGAGCTGGATTCCTTCAATTTCCATTTCACTC	1427
1381	QY	TCAGAAATAGATTGATTCATGAGATTAATGGCTCAGTTCTTTTGGAAAGTATATAAGGA	1440
1428	DB	TCAGAAATAGATTGATTCATGAGATTAATGGCTCAGTTCTTTTGGAAAGTATATAAGGA	1487
1441	QY	CGATCGCAAAATAAAATAGTGGCTATAAAAAGTTATCGAGCCAAATACCTACTGCTCCAAG	1500
1488	DB	CGATCGCAAAATAAAATAGTGGCTATAAAAAGTTATCGAGCCAAATACCTACTGCTCCAAG	1547
1501	QY	TCAGATATGAGATATGTTTTGCCGAGAGTGTCCATTTCTGCGAGCTCAATCATCCCTGC	1560
1548	DB	TCAGATATGAGATATGTTTTGCCGAGAGTGTCCATTTCTGCGAGCTCAATCATCCCTGC	1607
1561	QY	GTAATTCAGTTTGTGGGTGCTTGCTGAATGATCCAGCCAGTTTGCCATTTGCACTCAA	1620
1608	DB	GTAATTCAGTTTGTGGGTGCTTGCTGAATGATCCAGCCAGTTTGCCATTTGCACTCAA	1667
1621	QY	TACATATCAGGGGGTCTCTGTTCTCCCTTCATATGAGAGAGATTTCTGATTTG	1680
1668	DB	TACATATCAGGGGGTCTCTGTTCTCCCTTCATATGAGAGAGATTTCTGATTTG	1727
1681	QY	CAGTCTAAATTAATTATTCAGTATAGTTTGCCAAAGGATGGATGACTTCACAACTG	1740
1728	DB	CAGTCTAAATTAATTATTCAGTATAGTTTGCCAAAGGATGGATGACTTCACAACTG	1787
1741	QY	ACACAGCCAAATTATATCATCGTGAATTCGAAAGCTCAAAATTAATTTCTCTATGAGATGGG	1800

Db	1788	ACACAGCAATTATACATCGTGAATTTGAACAGTCACAATATTTCTTCTCTATGAGATGGG	1847
QY	1801	CATGCTGTGGTGGCAGATTTTGGAGAAATCAAGATTTTACAGTCTCTCGATGAAGACAAC	1860
Db	1848	CATGCTGTGGTGGCAGATTTTGGAGAAATCAAGATTTTACAGTCTCTCGATCAAGACAAC	1907
QY	1861	ATGCAAAACAAACCTGGGAACCTCCGTGGATGGCTCCTGAGTGTTCACGCAGTGCAC	1920
Db	1908	ATGCAAAACAAACCTGGGAACCTCCGTGGATGGCTCCTGAGTGTTCACGCAGTGCAC	1967
QY	1921	CGGTACACCAATCAAGAGCAGATGTCCTTCAGCTATGTCTGTCTGTCTGTGGAAATTTCTCACT	1980
Db	1968	CGGTACACCAATCAAGAGCAGATGTCCTTCAGCTATGTCTGTCTGTCTGTGGAAATTTCTCACT	2027
QY	1981	GGCGAAATTCATTTCCGTCTCATCTCAAGCAGCGGCTCGCGCAGCAGACATGGCTTACCAAC	2040
Db	2028	GGCGAAATTCATTTCCGTCTCATCTCAAGCAGCGGCTCGCGCAGCAGACATGGCTTACCAAC	2087
QY	2041	CACATCAGACCTCCCATTTGGCTATTTCATTTCCCAAGCCCATATCATCTCTGCTGATACGA	2100
Db	2088	CACATCAGACCTCCCATTTGGCTATTTCATTTCCCAAGCCCATATCATCTCTGCTGATACGA	2147
QY	2101	GGGTGGAAACGATGTCCTGAAGGAAGAACCGAATTTTCTGAAGTTGTTCATGAAGTTAGAA	2160
Db	2148	GGGTGGAAACGATGTCCTGAAGGAAGAACCGAATTTTCTGAAGTTGTTCATGAAGTTAGAA	2207
QY	2161	GAGTGTCTCTGCAACATTTGAGCTGATGTCCTCTGTCATCAAGTAACAGCAGTGGGTCTCTC	2220
Db	2208	GAGTGTCTCTGCAACATTTGAGCTGATGTCCTCTGTCATCAAGTAACAGCAGTGGGTCTCTC	2267
QY	2221	TCACCTCTCTTCTTCTGATTCGCTGGTGAACCGGGGAGGACCTTGGCGGAGTCATGTG	2280
Db	2268	TCACCTCTCTTCTTCTGATTCGCTGGTGAACCGGGGAGGACCTTGGCGGAGTCATGTG	2327
QY	2281	CGAGCATTAAGAAGTCGTTTCGAATTCGGAATATGCTCTAAATTCGAAGTCTCTATGCTGCT	2340
Db	2328	CGAGCATTAAGAAGTCGTTTCGAATTCGGAATATGCTCTAAATTCGAAGTCTCTATGCTGCT	2387
QY	2341	TTGTCCCAAGTGTGTGACAATATTTCTCTCAAGTGTCTCTTTTGGAGAGATGAAGA	2400
Db	2388	TTGTCCCAAGTGTGTGACAATATTTCTCTCAAGTGTCTCTTTTGGAGAGATGAAGA	2447
QY	2401	AGTCTTCAATACACACCCATTGACAAATATGGCTATGATTCGGATCCCATGAGCTCAATG	2460
Db	2448	AGTCTTCAATACACACCCATTGACAAATATGGCTATGATTCGGATCCCATGAGCTCAATG	2507
QY	2461	CATTTTCAATCTTTCGCCAAATPAGTAGCAGCTTTTGGAGACAGCAGC	2505
Db	2508	CATTTTCAATCTTTCGCCAAATPAGTAGCAGCTTTTGGAGACAGCAGC	2552

[illegible]

Matches 2505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	1	ATGGGAAATTTATAATCTAGACCAACCCAAACTTGTACTGATGAATGGAAAGGAAAAAGTC	60
Db	48	ATGGGAAATTTATAATCTAGACCAACCCAAACTTGTACTGATGAATGGAAAGGAAAAAGTC	107
Qy	61	AGTGAATCATATGTTATCAATAGAAAGATTAGAAAGATGACCTGCAGATCAAGGAAAAA	120
Db	108	AGTGAATCATATGTTATCAATAGAAAGATTAGAAAGATGACCTGCAGATCAAGGAAAAA	167
Qy	121	GAACCTGACGAACTTAAGGAATATATTTGGCTCTGATGAAGCCCTTCAGTAAAGTCAATTTA	180
Db	168	GAACCTGACGAACTTAAGGAATATATTTGGCTCTGATGAAGCCCTTCAGTAAAGTCAATTTA	227
Qy	181	AATTACCGCACTGAAATAGGCGTCTCTACTCTCATTTATGTTGCAATTTGTGGAGGCAAG	240
Db	228	AATTACCGCACTGAAATAGGCGTCTCTACTCTCATTTATGTTGCAATTTGTGGAGGCAAG	287
Qy	241	AAATCACATATTCGAACCTTATGTTGAAAGGGCTCCGCCATCTCGACTCACAGAAT	300
Db	288	AAATCACATATTCGAACCTTATGTTGAAAGGGCTCCGCCATCTCGACTCACAGAAT	347
Qy	301	GGATTTACAGCTTTGCAATTTAGCAGTTTACAAGGTAATGCAAGATTGATCACTTCTCTG	360
Db	348	GGATTTACAGCTTTGCAATTTAGCAGTTTACAAGGTAATGCAAGATTGATCACTTCTCTG	407
Qy	361	CTTCACAGTGGAGCTCATATACAGCAGGTTGGATACGGTGGCCCTCACTGCCCTCCATATT	420
Db	408	CTTCACAGTGGAGCTCATATACAGCAGGTTGGATACGGTGGCCCTCACTGCCCTCCATATT	467
Qy	421	GCTACAATAGCTGGCCACCTAGAGCGTCTGATGTGCTGTTGCCAACATGGAAGTAAATGTC	480
Db	468	GCTACAATAGCTGGCCACCTAGAGCGTCTGATGTGCTGTTGCCAACATGGAAGTAAATGTC	527
Qy	481	AATATTCAAGATGCGAGTTTTTTTCACTCAATGCAATATGCAAGGTAATGCAAGTAAATGTC	540
Db	528	AATATTCAAGATGCGAGTTTTTTTCACTCAATGCAATATGCAAGGTAATGCAAGTAAATGTC	587
Qy	541	CAGTAACTCGCCCTCTTTTGAATTTGGTGTGATGATGAAATGTAAGTGGTGAAGTTTGA	600
Db	588	CAGTAACTCGCCCTCTTTTGAATTTGGTGTGATGATGAAATGTAAGTGGTGAAGTTTGA	647
Qy	601	GATAGACCCCTCCACTAGACTCTGCAAAAGGATTTCTTGAATATGCAAACTCTTGATG	660
Db	648	GATAGACCCCTCCACTAGACTCTGCAAAAGGATTTCTTGAATATGCAAACTCTTGATG	707
Qy	661	GAAGAAGGAGCAAGACGATGTGAATGCTCAAGATATGAAGACCATGTCACCTCCAT	720
Db	708	GAAGAAGGAGCAAGACGATGTGAATGCTCAAGATATGAAGACCATGTCACCTCCAT	767
Qy	721	TTCTGTCTCGATTTGGACACCATGATAGTTAAGTATCTGCTGCAAGTGATTTGGAA	780
Db	768	TTCTGTCTCGATTTGGACACCATGATAGTTAAGTATCTGCTGCAAGTGATTTGGAA	827
Qy	781	GTTCACCTCATGTTGTTAATATCTAGAGTACCCCTTACACCTGGCATGCTCAAT	840
Db	828	GTTCACCTCATGTTGTTAATATCTAGAGTACCCCTTACACCTGGCATGCTCAAT	887
Qy	841	GGCAAAATTTGAAAGTTGCCAAGGAAATCATCAAAATATCAGGAAACAGAAAGTCTGACTAAG	900
Db	888	GGCAAAATTTGAAAGTTGCCAAGGAAATCATCAAAATATCAGGAAACAGAAAGTCTGACTAAG	947
Qy	901	GAAGACATCTTCAGTGAAGACGCTTTTCATAGTGGTTGTACCTATGCAAGACGATTTGAC	960
Db	948	GAAGACATCTTCAGTGAAGACGCTTTTCATAGTGGTTGTACCTATGCAAGACGATTTGAC	1007
Qy	961	CTAGTCAAAATTTCTTCTTGATCAGAAATGTCATAAATCAACCAAGGAAAGGATGGG	1020
Db	1008	CTAGTCAAAATTTCTTCTTGATCAGAAATGTCATAAATCAACCAAGGAAAGGATGGG	1067
Qy	1021	CACACTGGATTAACCTGCTGCTGCTACACCGGTCAATTCGCTGGTTTCAAGTTCTTACTG	1080
Db	1068	CACACTGGATTAACCTGCTGCTGCTGCTACACCGGTCAATTCGCTGGTTTCAAGTTCTTACTG	1127

QY 2161 GAGTGTCTCTGCAACATTGAGCTGATGTCCTCTGATCAAGTAACAGAGTGGTCTCTC 2220  
DB 2208 GAGTGTCTCTGCAACATTGAGCTGATGTCCTCTGATCAAGTAACAGAGTGGTCTCTC 2267  
QY 2221 TCACCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2280  
DB 2268 TCACCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2327  
QY 2281 GCAGCATTAAGAAGTCGTTTTCGAATTGGAATATGCTCTAAATGCAAGTCTCTATGCTGCT 2340  
DB 2328 GCAGCATTAAGAAGTCGTTTTCGAATTGGAATATGCTCTAAATGCAAGTCTCTATGCTGCT 2387  
QY 2341 TTGTCCCAAGTCTGGACAAATATTCCTCTCAAGTCTGTCCTTTGGAGGAGATGAAGA 2400  
DB 2388 TTGTCCCAAGTCTGGACAAATATTCCTCTCAAGTCTGTCCTTTGGAGGAGATGAAGA 2447  
QY 2401 AGTCTTCAATACACACCCCATTCACAAATATGGCTATGATCCGATCCCATGAGCTCAATG 2460  
DB 2448 AGTCTTCAATACACACCCCATTCACAAATATGGCTATGATCCGATCCCATGAGCTCAATG 2507  
QY 2461 CATTTTCATCTTGGCGAAATAGTAGAGCTTTGAGGACAGCAGC 2505  
DB 2508 CATTTTCATCTTGGCGAAATAGTAGAGCTTTGAGGACAGCAGC 2552

RESULT 9  
AR437545  
LOCUS AR437545 3025 bp DNA linear PAT 18-DEC-2003  
DEFINITION Sequence 1 from patent US 6660490.  
AR437545  
AR437545.1 GI:40202640  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 3025)  
AUTHORS Raju, J.  
TITLE CARP protein and nucleic acid molecules and uses therefor  
JOURNAL Patent: US 6660490-A 1 09-DEC-2003;  
FEATURES Location/Qualifiers  
source  
1..3025  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN  
Query Match 100.0%; Score 2505; DB 6; Length 3025;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGGGAAATATAAATCTAGACCAACCCAACTTGTACTGATGAATGGAAGAAAAAGTC 60  
DB 48 ATGGGAAATATAAATCTAGACCAACCCAACTTGTACTGATGAATGGAAGAAAAAGTC 107  
QY 61 AGTGAATCATATGTTATCAATAGAAAGATTAGAGATGACCTGCAGATCAAGGAAAA 120  
DB 108 AGTGAATCATATGTTATCAATAGAAAGATTAGAAAGATGACCTGCAGATCAAGGAAAA 167  
QY 121 GAACTGCAGAACTAAGGAATATTTGGCTCTGATGAAGCCCTTCAGTAAAGTCAATTTA 180  
DB 168 GAACTGCAGAACTAAGGAATATTTGGCTCTGATGAAGCCCTTCAGTAAAGTCAATTTA 227  
QY 181 AATTACCGACTGAAATGGGCTGCTCTACTTTCATTTTANGTTGATTTGGAGGCAAG 240  
DB 228 AATTACCGACTGAAATGGGCTGCTCTACTTTCATTTTANGTTGATTTGGAGGCAAG 287  
QY 241 AATTCATATTCGAACTCTTATGTTGAAGGGCTCCGCCCATCTCGACTGACAGAATAAT 300  
DB 288 AATTCATATTCGAACTCTTATGTTGAAGGGCTCCGCCCATCTCGACTGACAGAATAAT 347  
QY 301 GGATTTACAGCCTTGCATTTAGCAGTTTACAAGGATTAATGCAGAAATGATCACTTCTCTG 360  
DB 348 GGATTTACAGCCTTGCATTTAGCAGTTTACAAGGATTAATGCAGAAATGATCACTTCTCTG 407

QY 361 CTTTCACTGGAGCTGATATACAGCAGGTTGGATACCGTGGCCTCACTGCCCTCCCATATT 420  
DB 408 CTTTCACTGGAGCTGATATACAGCAGGTTGGATACCGTGGCCTCACTGCCCTCCCATATT 467  
QY 421 GCTCAATAGCTGGCCACCTAGAGGCTGCTGATGTCGTGTTGTCACCAATGGAGCTAATGTC 480  
DB 468 GCTCAATAGCTGGCCACCTAGAGGCTGCTGATGTCGTGTTGTCACCAATGGAGCTAATGTC 527  
QY 481 AATATTCAAGATGCAAGTCTTTTTCACCTCCATGTCATATTGTCAGGCTACTATGACATGAA 540  
DB 528 AATATTCAAGATGCAAGTCTTTTTCACCTCCATGTCATATTGTCAGGCTACTATGTCATGAA 587  
QY 541 CAGGTAACCTCGCCTCTCTTTTGAATTTGGTCTCATGTAAATGTAAAGTGGTGAAGTTGGA 600  
DB 588 CAGGTAACCTCGCCTCTCTTTTGAATTTGGTCTCATGTAAATGTAAAGTGGTGAAGTTGGA 647  
QY 601 GATAGACCCCTCCACCTAGCATCTGCAAAAGGATTTCTTGAATATTGCAAACTCTTGATG 660  
DB 648 GATAGACCCCTCCACCTAGCATCTGCAAAAGGATTTCTTGAATATTGCAAACTCTTGATG 707  
QY 661 GAAGAAGCGCAGCAAGCAGATGTAAGTCTCAAGATAATGAAGACCATGTCCTCACTCCAT 720  
DB 708 GAAGAAGCGCAGCAAGCAGATGTAAGTCTCAAGATAATGAAGACCATGTCCTCACTCCAT 767  
QY 721 TTCTGTCTCTCGATTTGGACACCATGATATAGTAAAGTATCTGCTGCAAAAGTATTCGAA 780  
DB 768 TTCTGTCTCTCGATTTGGACACCATGATATAGTAAAGTATCTGCTGCAAAAGTATTCGAA 827  
QY 781 GTTCAACCTCATGTTGTTAAATATCTATGGAGATACCCCTTACACCTGGCATGCTCAAT 840  
DB 828 GTTCAACCTCATGTTGTTAAATATCTATGGAGATACCCCTTACACCTGGCATGCTCAAT 887  
QY 841 GGCAAAATTTGAAGTTGCCAAGGAAATCATCCAAATATCAGGAAACAGAAAGTCTGACTAAG 900  
DB 888 GGCAAAATTTGAAGTTGCCAAGGAAATCATCCAAATATCAGGAAACAGAAAGTCTGACTAAG 947  
QY 901 GAAACATCTTCAGTGAACACAGCTTTTCATAGTGGTGTGACTATGTCAGCAAGCAATGAC 960  
DB 948 GAAACATCTTCAGTGAACACAGCTTTTCATAGTGGTGTGACTATGTCAGCAAGCAATGAC 1007  
QY 961 CTAGTCAAAATTTCTTGTATCAGAACTGTCATAAATCATCAACCAAGGAAAGGATGGG 1020  
DB 1008 CTAGTCAAAATTTCTTGTATCAGAAATGTCAATCAATCAACCAAGGAAAGGATGGG 1067  
QY 1021 CACACTGGATTAACCTCTGCTTGTACACCGTCAATTCGCTGGTTCAGTCTTCTACTG 1080  
DB 1068 CACACTGGATTAACCTCTGCTTGTACACCGTCAATTCGCTGGTTCAGTCTTCTACTG 1127  
QY 1081 GATAATGGAGCTGATGATGAATCTAGTGGCTTGTGATCCCGCAGGCTCTAGTGGTGAATA 1140  
DB 1128 GATAATGGAGCTGATGATGAATCTAGTGGCTTGTGATCCCGCAGGCTCTAGTGGTGAATA 1187  
QY 1141 GATGACGACATGTTTGTATGTTGGCTTATGAAAAGGGCATGATGCCATTTGTCACACTC 1200  
DB 1188 GATGACGACATGTTTGTATGTTGGCTTATGAAAAGGGCATGATGCCATTTGTCACACTC 1247  
QY 1201 CTGAACCATTTAAGAGACCAAGATGAATTCGCTGTATGAATATTTCTCAGCTTGA 1260  
DB 1248 CTGAACCATTTAAGAGACCAAGATGAATTCGCTGTATGAATATTTCTCAGCTTGA 1307  
QY 1261 GGAGATGGCTCTCTATGTCCTTCCATCACTCCCTTGGGAAAGATTAAAGCATGACAAAA 1320  
DB 1308 GGAGATGGCTCTCTATGTCCTTCCATCACTCCCTTGGGAAAGATTAAAGCATGACAAAA 1367  
QY 1321 GAGAAGCGCAGATATTTCTCTCTTAAGAGCTGATGTCCTTCAATTTCCATCTTCAGCTC 1380  
DB 1368 GAGAAGCGCAGATATTTCTCTCTTAAGAGCTGATGTCCTTCAATTTCCATCTTCAGCTC 1427  
QY 1381 TCAGAAATTTGAGTTCCATGAGATTTATGCTCAGGCTCTTTTGGGAAAGTATATAAGGA 1440  
DB 1428 TCAGAAATTTGAGTTCCATGAGATTTATGCTCAGGCTCTTTTGGGAAAGTATATAAGGA 1487  
QY 1441 CGATGCAAGAAATAAATAATAGTGGCTATAAAAACTTTATGAGGCAATACCTACTGCTCCAAG 1500

Db CGATGCGAGAAATAAATAGTGGCTATATAACGGTATATGAGCCAAATACCTACTGCTCCAAG 1547  
Qy TCAGATGTGGATATGTTTGGCGAGAGGTGTCCATCTCTGCCAGGTCAATCATCCTCCG 1560  
Db TCAGATGTGGATATGTTTGGCGAGAGGTGTCCATCTCTGCCAGGTCAATCATCCTCCG 1607  
Qy GTAATTCAGTTTGGTGGTGTCTGTTGAATGATCCAGCCAGTTGCCATTTGCACTCAA 1620  
Db GTAATTCAGTTTGGTGGTGTCTGTTGAATGATCCAGCCAGTTGCCATTTGCACTCAA 1667  
Qy TACATATCAGGGGTTCTCTGTTCTCTCTCTCATGAGCAGAGAGGATCTCTGATTTG 1680  
Db TACATATCAGGGGTTCTCTGTTCTCTCTCTCATGAGCAGAGAGGATCTCTGATTTG 1727  
Qy CAGTCTAAATTAATTTATGTCAGTATGTTGCCAAAGGCAATGAGTACCTTCAACCTG 1740  
Db CAGTCTAAATTAATTTATGTCAGTATGTTGCCAAAGGCAATGAGTACCTTCAACCTG 1787  
Qy ACAGAGCAATATACATCTGTCATCTGAAAGTCAATATCTCTCTCATGAGGATGGG 1800  
Db ACAGAGCAATATACATCTGTCATCTGAAAGTCAATATCTCTCTCATGAGGATGGG 1847  
Qy CATGCTGTGGTGGCAGATTTTGGAGATCAAGATTTCTACAGTCTCTGGATGAAGACAAC 1860  
Db CATGCTGTGGTGGCAGATTTTGGAGATCAAGATTTCTACAGTCTCTGGATGAAGACAAC 1907  
Qy ATGACAAAACAACTGGGAACCTCCGTTGGATGGCTCTGAGTGTTCACGAGTGCCT 1920  
Db ATGACAAAACAACTGGGAACCTCCGTTGGATGGCTCTGAGTGTTCACGAGTGCCT 1967  
Qy CGGTACACCATCAAGCAGATGCTTCAGTATGCTCTGTGTCGTGGCAATTTCTCACT 1980  
Db CGGTACACCATCAAGCAGATGCTTCAGTATGCTCTGTGTCGTGGCAATTTCTCACT 2027  
Qy GGCAGAAATTCATTCGTCATCTCAAGCCAGCGGTCCGCGACGACATGCTTTACCA 2040  
Db GGCAGAAATTCATTCGTCATCTCAAGCCAGCGGTCCGCGACGACATGCTTTACCA 2087  
Qy CACATCAGCTCCCATTTGCTATTCATTTCCCAAGCCATATCATCTCTGTGATACGA 2100  
Db CACATCAGCTCCCATTTGCTATTCATTTCCCAAGCCATATCATCTCTGTGATACGA 2147  
Qy GGTGGAAACGATGCTCTGAAGGAAGACCCGAAATTTCTCAAGTGTCTCATGAAGTAGAA 2160  
Db GGTGGAAACGATGCTCTGAAGGAAGACCCGAAATTTCTCAAGTGTCTCATGAAGTAGAA 2207  
Qy GAGTGTCTTCTGCAACATGAGCTGATGTCCTGTCATCAAGTAAACGAGTGGTCTCTC 2220  
Db GAGTGTCTTCTGCAACATGAGCTGATGTCCTGTCATCAAGTAAACGAGTGGTCTCTC 2267  
Qy TCACCTTCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2280  
Db TCACCTTCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2327  
Qy GCAGCATTAAGAAATGCTTTGGAATGCTTCTAAATGCAAGTCTTATGCTGCT 2340  
Db GCAGCATTAAGAAATGCTTTGGAATGCTTCTAAATGCAAGTCTTATGCTGCT 2387  
Qy TTGTCCTCAAGTGTGCAACATATCTCTCAAGTGTCTCTTTTGGAGGAGTAAAGA 2400  
Db TTGTCCTCAAGTGTGCAACATATCTCTCAAGTGTCTCTTTTGGAGGAGTAAAGA 2447  
Qy AGTCTTCAATACACACCATTTGCAAAATATGCTATGATCCGATCCCATGAGTCAATG 2460  
Db AGTCTTCAATACACACCATTTGCAAAATATGCTATGATCCGATCCCATGAGTCAATG 2507  
Qy CATTTTCATTTCTGCGAATATGACGCTTTGAGGACAGCAGC 2505  
Db CATTTTCATTTCTGCGAATATGACGCTTTGAGGACAGCAGC 2552

BD269562  
LOCUS Novel CARP protein and nucleic acid molecules and uses therefor.  
DEFINITION Novel CARP protein and nucleic acid molecules and uses therefor.  
ACCESSION BD269562  
VERSION BD269562.1 GI:33079330  
KEYWORDS JP 2002535962-A/4.  
SOURCE Rattus norvegicus (Norway rat)  
ORGANISM Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
REFERENCE 1 (bases 1 to 2505)  
AUTHORS Raju,J.  
TITLE Novel CARP protein and nucleic acid molecules and uses therefor  
JOURNAL Patent: JP 2002535962-A 4 29-OCT-2002;  
MILLENNIUM PHARMACEUTICALS INC  
COMMENT OS Rattus norvegicus (rat)  
PN JP 2002535962-A/4  
PD 29-OCT-2002  
PF 10-DEC-1999 JP 2000586772  
PR 11-DEC-1998 US 60/111938,14-APR-1999 US 09/291839 PI  
JEYASEELAN RAJU  
PC C12N15/09,C07K16/40,C12N1/15,C12N1/19,C12N1/21,C12N5/10,C12N9/  
PC 10,C12Q1/02,  
PC C12Q1/68,G01N33/15,G01N33/50,G01N33/53,G01N33/68//A61K38/45,  
PC A61K39/395,  
PC A61K45/00,A61P9/04,A61P9/10,A61P35/00,A61P43/00,C12N15/00, PC  
C12N5/00,  
PC A61K37/52  
CC Novel CARP protein and nucleic acid molecules and uses CC  
therefor  
FH Key Location/Qualifiers  
FT CDS Location/Qualifiers  
source 1..2505  
/organism="Rattus norvegicus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10116"  
ORIGIN  
Query Match 75.5%; Score 1892.2; DB 6; Length 2505;  
Best Local Similarity 84.7%; Pred. No. 0;  
Matches 2122; Conservative 0; Mismatches 383; Indels 0; Gaps 0;  
Qy 1 ATGGGAAATTAATAATCTAGACCAACCCAACTTCTACTGATGAATGGAAGAAAAAGTC 60  
Db 1 ATGGGAAATTAATAATCTAGACCAACCAACAGACTTCTCTGATGATGGAAGAAAGATT 60  
Qy 61 AGTGAATCATATGTTATCATCAATAGAAAGATTAGAAAGATGACCTGCGAGATCAAGAAAA 120  
Db 61 AGTGAATCTTACGCTATTATCATAGAAAGGCTGGAGGATAACCTGCGAGATCAAGAAAA 120  
Qy 121 GAAGTACAGAACTAAGGAATATATTGCTCTGATGAAGCTTCAGTAAAGTCAATTTA 180  
Db 121 GAATTTCAAGAACTAAGGCAATCTTTGGCTCTGATGAAGCTTCAGTAAAGTCAATTTA 180  
Qy 181 AATTACCGCACTGAAATGGGCTGCTCTACTTCAATTTATTTGTTGTTGTTGTTGTTGTTG 240  
Db 181 AATTACCGCACAGAGCGTGGCTGCTCTGCTTACACCTCTGCTGCTGCTGCTGCTGCTGCTG 240  
Qy 241 AATACATATTTGCAACTTATGTTGAAGGGTCCGCCCTCCGCTCCGCTCCGCTCCGCTCCG 300  
Db 241 AAGTCAATATCCGTGCGCTTATGTTAAAGGGTCCCGTCCATCCAGCTGACGAGAAAT 300  
Qy 301 GGATTTACAGCTTGCATTTAGCAGTTTACAGGATATGCAAGTATGATGATGATGATGATG 360  
Db 301 GGATTTACAGCTTGCAGCTGCGCTGCTTTTACAAGGACAGCCGCGAACTTATCCTTCACTG 360  
Qy 361 CTTTCAGTGGAGCTGATATACAGCAGTTTGGATACGTTGGTGGTGGTGGTGGTGGTGGTGG 420  
Db 361 TTGTCAGCGGAGCAGATGTTTACAGAGTGGATACGTTGGTGGTGGTGGTGGTGGTGGTGG 420  
Qy 421 GCTACAATAGCTGGCCACTAGAGGCTGCTGATGCTGTTGTTGCAACATGAGCTTAATGTC 480

Db 421 GCTGCNATAGCTGGACACCCAGAGGCTCGAGAGTGTCTGTCTACAACTATGGGCCCAATGTG 480  
QY 481 AATATTCAAGATGCAAGTCTTTTCTACTCCATTCGCAATATGAGCGTACTATGAGCATGAA 540  
Db 481 AATGTTCAAGATGCGGTCTTCTTCCACCCACTGACATTCGAGGCTACTATGGGCACGAG 540  
QY 541 CAGGTAACTCGCCTTCTTTTGAATTTTGGTCTGATGTAATGTAAGTGGTGAAGTTGGA 600  
Db 541 CAGGTAACTCGCCTTCTTTTGAATTTTGGTCTGATGTAATGTAAGTGGTGAAGTTGGA 600  
QY 601 GATAGACCCCTCCACTAGCATCTGCAAAAGAGATCTTTGAATATTTGCAAACTCTTGATG 660  
Db 601 GACAGGCTCTGCACCTGCGCTCTGCAAAAGGCTCTTTCAACATTTGAAACTCTCTGTA 660  
QY 661 GAAGAAGGAGCAAAAGCAGATGTAATCTCAAGTATATGAGCCATGTCCTCACTCCAT 720  
Db 661 GAAGAAGGAGCAAAAGCAGATGTAATCTCAAGTATATGAGCCATGTCCTCACTCCAT 720  
QY 721 TTCTGTTCTCGATTTGGACACCATGATATAGTAAAGTATCTGCTGCAAAAGTGAATGGAA 780  
Db 721 TTCTGTTCTCGATTTGGACACCATGATATAGTAAAGTATCTGCTGCAAAAGTGAATGGAA 780  
QY 781 GTTCAACCTCATGTTGTTAAATCTATGAGATACCCCTTTACCTGGCATGTCTACAAT 840  
Db 781 GTTCAACCTCATGTTGTTAAATCTATGAGATACCCCTTTACCTGGCATGTCTACAAT 840  
QY 841 GGCATAATTTGAAGTTCGCAAGGAATCATCCAAATATCAGGAACAGAAAGTCTGACTAAG 900  
Db 841 GGCATAATTTGAAGTTCGCAAGGAATCATCCAAATATCAGGAACAGAAAGTCTGACTAAG 900  
QY 901 GAAAACATCTTCAGTGAACACAGCTTTTCATAGTGTGTGATGCTGATGGAAGAGATGAC 960  
Db 901 GAAAACATCTTCAGTGAACACAGCTTTTCATAGTGTGTGATGCTGATGGAAGAGATGAC 960  
QY 961 CTAGTCAAAATTTCTTCTGATCAGATGTCATAACATCAACCAACCAAGGAGGATCGG 1020  
Db 961 CTAGTCAAAATTTCTTCTGATCAGATGTCATAACATCAACCAACCAAGGAGGATCGG 1020  
QY 1021 CACACTGGATTTACACTCTGCTGCTACCAAGTGCATCTCGCTGGTTCAGTTCCTACTG 1080  
Db 1021 CACACTGGATTTACACTCTGCTGCTACCAAGTGCATCTCGCTGGTTCAGTTCCTACTG 1080  
QY 1081 GATAATGGAGCTGATATGATATGATGCTGTGATGCCAGAGCTCTAGTGGTGAATAA 1140  
Db 1081 GATAATGGAGCTGATATGATATGATGCTGTGATGCCAGAGCTCTAGTGGTGAATAA 1140  
QY 1141 GATGAGCAGACATGTTGATGTTGGCTTATGAAAAGGCGCATGATGCCATTTGTCACATC 1200  
Db 1141 GATGAGCAGACATGTTGATGTTGGCTTATGAAAAGGCGCATGATGCCATTTGTCACATC 1200  
QY 1201 CTGAAGCATTTAAGAGACACCAAGATGAATGGCCCTGTAATGAATATTTCTCAGCCTGGA 1260  
Db 1201 CTGAAGCATTTAAGAGACACCAAGATGAATGGCCCTGTAATGAATATTTCTCAGCCTGGA 1260  
QY 1261 GGAGATGGCTCTGATGCTGTGCTTCCATCACCTTGGGAGATTAAGAGCATGACAAAA 1320  
Db 1261 GGAGATGGCTCTGATGCTGTGCTTCCATCACCTTGGGAGATTAAGAGCATGACAAAA 1320  
QY 1321 GAGAAGGAGCATATTTCTCTCTTAAGAGCTGGAATTCCTTCAATTTCCATTTCCAGTCT 1380  
Db 1321 GAGAAGGAGCATATTTCTCTCTTAAGAGCTGGAATTCCTTCAATTTCCATTTCCAGTCT 1380  
QY 1381 TCAGAAATGAGTTCATGATGATGATGCTAGTTCCTTTGGGAAAGTATTAAGGA 1440  
Db 1381 TCAGAAATGAGTTCATGATGATGATGCTAGTTCCTTTGGGAAAGTATTAAGGA 1440  
QY 1441 CGATGCGAATAAATAAGTGGCGATCAACAGATACCGAGCCACACCTACTGCTCCAG 1500  
Db 1441 CGATGCGAATAAATAAGTGGCGATCAACAGATACCGAGCCACACCTACTGCTCCAG 1500  
QY 1501 TCAGATGTGATATGTTTTGCCAGAGGTCTCCATTTCTGTCAGGTCTCAATCATCTCTGC 1560

RESULT 11  
AR269114  
LOCUS  
DEFINITION  
ACCESSION  
VERSION

AR269114  
Sequence 9 from patent US 6500654.  
AR269114  
AR269114.1

2505 bp

DNA

linear

PAT 10-APR-2003

GI:29699957

Db 1501 TCAGAGCTGGATATGTTTTTCCGAGAGGTGTCCATTTCTGCCAGCTCAACACCCCTGC 1560  
QY 1561 GTAAATTCAGTTTGGGTGCTTGTGTAATGATCCAGCCAGTTTGGCATTGTCACTCAA 1620  
Db 1561 GTGGTTTCAGTTTGGGTGCTTGTGTAATGATCCAGCCAGTCAATTTGCCATTGTCACTCAG 1620  
QY 1621 TACATATCAGGGGGTCTCTGTTCTCCCTCTTCATGAGCAGAGAGATTTCTTGATTTG 1680  
Db 1621 TACATATCAGGGGGTCTCTGTTCTCCCTCTTCATGAGCAGAGAGATTTCTTGATTTG 1680  
QY 1681 CAGTCTAAATTAATTTATGTCAGTAGATGTTCCCAAGGCGATGGAGTACCTTCAACAACCTG 1740  
Db 1681 CAGTCTAAATTAATTTATGTCAGTAGATGTTCCCAAGGCGATGGAGTACCTTCAACAACCTG 1740  
QY 1741 ACACAGCCAAATTTATACATCTGACTTGAACAGTCAACAATTTCTTCTATGAGAGATGGG 1800  
Db 1741 ACCAGCCAAATCATACACCGGACCTGAAACAGCCACAATATTTCTGCTCTATGAGGATGGC 1800  
QY 1801 CATGCTGTGGTGGCAGATTTTGGAGATCAAGATTTCTACAGTCTCTGGATGAAGACAAC 1860  
Db 1801 CATGCTGTGGTGGCAGATTTTGGAGATCAAGATTTCTACAGTCTCTGGATGAAGACAAC 1860  
QY 1861 ATGACAAAACAACTTGGGAACTCCGTTGGATGGCTCCTGAGGTGTTTCAACAGTGCAC 1920  
Db 1861 ATGACAAAACAACTTGGGAACTCCGTTGGATGGCTCCTGAGGTGTTTCAACAGTGCAC 1920  
QY 1921 CGGTACACCATCAAAAGCAGATGCTTTACAGTATGCTGTGCTGTCTGTGGGAAATTTCTCACT 1980  
Db 1921 AGATACACCATCAAGGCTGATGCTTTCAGTTACTCCCTGTCTGTGGGAGTCTCTCACT 1980  
QY 1981 GCGGAAATTTCCATTCGCTCATCTCAAGCCAGCGGCTGCGGAGCAGAGATGGCTTACCAC 2040  
Db 1981 GCGGAAATTTCCATTCGCTCATCTCAAGCCAGCGGCTGCGGAGCAGAGATGGCTTACCAC 2040  
QY 2041 CACATCAGACCTCCATTTGGCTTATTCATTTCCAAAGCCCATATCATCTCTCTGATAGGA 2100  
Db 2041 CACATCAGACCTCCATTTGGCTTATTCATTTCCAAAGCCCATATCATCTCTCTGATAGGA 2100  
QY 2101 GGGTGAACCGATGCTCTGAAAGAGACCCGNAATTTCTGAAGTTGTCATGAAGTTAGAA 2160  
Db 2101 GGGTGAACCGATGCTCTGAAAGAGACCCGNAATTTCTGAAGTTGTCATGAAGTTAGAA 2160  
QY 2161 GAGTGTCTCTGCAACATTTGAGCTGATGCTCTCTGATCAAGTAAACAGAGTGGTCTCTC 2220  
Db 2161 GAGTGTCTCTGCAACATTTGAGCTGATGCTCTCTGATCAAGTAAACAGAGTGGTCTCTC 2220  
QY 2221 TCACCTTCTTCTTCTGATTTGCTGTAACCGGAGGAGCCTGCGGAGGATCATGTG 2280  
Db 2221 TCACCTTCTTCTTCTGATTTGCTGTAACCGGAGGAGCCTGCGGAGGATCATGTG 2280  
QY 2281 GCAGCATTAAGAGTCTGTTTTCGAATTTGAATATGCTCTAAATGCAAGGTCTCTATGCTGCT 2340  
Db 2281 GCAGCATTAAGAGTCTGTTTTCGAATTTGAATATGCTCTAAATGCAAGGTCTCTATGCTGCT 2340  
QY 2341 TTCTCCCAAGTCTGGAACAATTTCTCTCTCAAGTGTCTCTTTGGAGAGATGAAGA 2400  
Db 2341 TTCTCCCAAGTCTGGAACAATTTCTCTCTCAAGTGTCTCTTTGGAGAGATGAAGA 2400  
QY 2401 AGTCTTCAATACACCCCATTTGACAAATATGCTATGATCCGATCCCATGAGTCAATG 2460  
Db 2401 AGTCTTCAATACACCCCATTTGACAAATATGCTATGATCCGATCCCATGAGTCAATG 2460  
QY 2461 CATTTTCAATTTGCGGAAATATGATGACGCTTTGAGGACAGCAGC 2505  
Db 2461 CATTTTCAATTTGCGGAAATATGATGACGCTTTGAGGACAGCAGC 2505

KEYWORDS	Unknown.
SOURCE	Unknown.
ORGANISM	Unclassified.
REFERENCE	1 (bases 1 to 2505)
AUTHORS	Raju, J.
TITLE	CARK protein and nucleic acid molecules and uses therefor
JOURNAL	Patent: US 6500654-A 9 31-DEC-2002;
FEATURES	Location/Qualifiers
source	1..2505
/organism="unknown"	
/mol_type="genomic DNA"	
ORIGIN	
Query Match	75.5%; Score 1892.2; DB 6; Length 2505;
Best Local Similarity	84.7%; Pred. No. 0;
Matches 2122; Conservative	0; Mismatches 393; Indels 0; Gaps 0;
QY	1 ATGGGAATTATAATCTAGACCAACCCAACTTGTACTGATGAATGGAGAAAAGTC 60
Db	1 ATGGGAATTACAAATCCAGACCAACACAGACTTGTCTCATGAATGGAGAAAAGTT 60
QY	61 AGTGATCATATGTATACAAATAGAAAGATTAGAAGATGACCTGCAGATCAAGSAAAA 120
Db	61 AGTGATCTTACGCTATTATCATAGAAAGGCTGGAGGATAACCTGCAGATCAAGSAAAA 120
QY	121 GAACTGCAGAACTAAGGAATATATTTGGCTCTGATGAAGCCTTCAGTAAAGTCAATTTA 180
Db	121 GAAATTCAGAACTAAGGCACATCTTTGGCTCTGATGAAGCCTTCAGTGAAGTCAATTTA 180
QY	181 AATTACCGCATGAAATGGCTGTCTCTACTTCAATTTATGTTGCAATTTGTGGAGGCAAG 240
Db	181 AATTACCGCACAGAGCGTGGCTGTCTCTCTACACCTCTGTGTCTGTGGGCGCAAC 240
QY	241 AAATCACATATTCGAATCTTATGTTGAAAGGCTCGCCCATCTCGACTGACAGAAAT 300
Db	241 AAGTCANATTCGTGCGCTTATGTTAAAGGCTCGTCCATCCAGATGACGAGAAAT 300
QY	301 GGATTTACGCTTTCATTTAGCAGTTTACAGAGTAATGACAGAAATTTGATCACTTCTCTG 360
Db	301 GGGTTTCAGCTCTGCACCTGGCGGTTTACAGAGACAGCCGGAACCTTATCACTTCACTG 360
QY	361 CTTACAGTGGAGTGATACAGAGTTGATAGCGTGGCTCACTGCGCTCCATAT 420
Db	361 TTGCACAGCGAGCAGATGTTACAGAGTGGGATACGGTGGCTTCAAGCCCTCCACATA 420
QY	421 GCTCAATAGCTGCCACCTAGAGGCTGTGATGTGCTGTGCAACATGAGACTAATGTC 480
Db	421 GCTGCAATAGCTGGACACCCAGAGGCTGCAGAGTGTCTGTACAACTGGGCGCAATGTG 480
QY	481 AATATTCAGATGAGTATACAGAGTTGATAGCGTGGCTCACTGCGCTCCATAT 540
Db	481 AATGTTCAAGATGCGGTCCTTTCACCCCACTGCACATTCGAGCCTTACTATGGCAGAG 540
QY	541 CAGGTAACCTGCGCTTCTTTGAAATTTGGTGTGATGTAAATGTAAGTGGTGAAGTTGGA 600
Db	541 CAGGTAACCAAGTGTCTTTTGAAGTTTGGTGTGATGTCAATGTAAGCGGTGAAGTTGG 600
QY	601 GATGAGCCCTCCACCTAGCATCTGCAAAAGGATTTCTGAAATTTGCAAACTCTTGATG 660
Db	601 GACAGGCTCTGCACCTGGCTCTGCAAAAGGCTTCTTCAACATTTGTGAACTCTCGTA 660
QY	661 GAAGAGGCGAGCAAGCAGATGTGAATGTCAAGATAATGAAGACCATGTCCCATCCAT 720
Db	661 GAAGAGGAGCAAGCAGATGTGAACGTTCAGGACATGAAGACCACTGCTCCCTGAC 720
QY	721 TTCTGTTCTCGATTTGACACCATGATATAGTATGATCTGCTGCAAGTGAATTTGAA 780
Db	721 TTCTGTTCTCGATTTGACACCATGATATAGTATGATCTGCTGCAAGTGAATTTGAA 780
QY	781 GTTCAACCTCATGTTGTTATATCTATGGAGATACCCCTTACCTGCGATGCTACAAAT 840
Db	781 GTCCAGCTCACGTCTAATCACTATGTTGACATCTCTTTGCACTGCTGGCATGCTACANT 840
QY	841 GGCAAAATTTGAAGTTGCCAAGGAAATCATCCAAATATCAGGAACAGAAAGTCTGACTAAG 900
Db	841 GGAAATTTTGAAGTTGCCAAGGAAATTTGTCAGGTAAACAGGAATCTGAAAGTCTGACTAAG 900
QY	901 GAAACATCTTTCAAGTGAACAGCTTTTTCATAGTGTCTGTACCTATGGCAAGAGCATTTGAC 960
Db	901 GAAACATCTTTCAAGGAGACAGCTTTTTCAGAGTCTGTACCTATGGCAAGAGCATTTGAC 960
QY	961 CTAGTCAAAATTTCTTTGATCAGAATGTCAATAAATCAACCAACCAAGAGAGGATGGG 1020
Db	961 CTGTCAAAATTTCTTTGATCAGAATGTCTGTGAACATTAACCAACCGAGAGAGATGGG 1020
QY	1021 CACACTGATTACACTCTGCTTACACAGGTCATTCATTCGCTCGCTGTTTTCAGTTCCTTACTG 1080
Db	1021 CACAGGATTCACCTCTGCTTCTACACAGGCTATCCGCTGTTTTCAGTTCCTTACTT 1080
QY	1081 GATAATGGAGCTGATATGAATCTAGTGGCTTGTGATCCAGCAGGTCTAGTGGTGA AAAA 1140
Db	1081 GATAATGGTGCAGATATGAATCTTGTGCTTGTGATCCAGCAGGTCTAGTGGTGA AAAA 1140
QY	1141 GATGAGCAGACATGTTTGTATGTTGGCTTATGAAAGGCGCATGATGCCATTGTCACTC 1200
Db	1141 GATGAGCAGACATGTTTGTATGTTGGCTTACGAGAAAGGACATGATGCCATTGTACACTC 1200
QY	1201 CTGAAGCATTTAAGAGACCAAGATGAATTTGCCCTGTATGAATATTTCTCAGCCTGGA 1260
Db	1201 CTGAAGCCTTACAAGAGACCCAGGAGGAGTGCCTATTAAGATATTTCCAGCCTGGA 1260
QY	1261 GGAGATGGCTCTATGTGTCTGTTCATCACCTTTGGGGAGAGATTAAAGCATGACAAAA 1320
Db	1261 GGAGATGGCTCTATGTGTCTGTTCCTTCCCTTTGGGCAAGATTAAAGCATGACAAAA 1320
QY	1321 GAGAAGCAGATATTTCTCTTAAGAGCTGATTTGCCCTTCACTTTCCATTTCTCAGCTC 1380
Db	1321 GAGAAGCAGATGTTCTCTCTGAGGCTGAATCTACCTCCCGCTTCATCTCCAATC 1380
QY	1381 TCAGAAATTTGAGTTCCATGAGATTTTGGCTCAGGTTCTTTTGGGAAAGTATATAAGGA 1440
Db	1381 TCCGAAATCGAGTTCCACGAGATTTACGGCTCGGTTCTTTTGGGAAAGTCTATAAGGG 1440
QY	1441 CGATGAGAAATATAAATAGTGGCTATAAAGCTTATCGAGCCATACCTACTGCTCCAAG 1500
Db	1441 CGATGAGAAATATAAATAGTGGCTATAAAGCTTATCGAGCCATACCTACTGCTCCAAG 1500
QY	1501 TCAGATGTGGATATGTTTTCGCGAGAGGTGTCATTTCTGCGAGCTCAATCATCCCTGTC 1560
Db	1501 TCAGAGCTGGATATGTTTTCGCGAGAGGTGTCATTTCTGCGAGCTCAACCACTGTC 1560
QY	1561 GTAAATTCAGTTTGTGGTGTCTGCTGATGATGATCCAGCAGTTTGCCATTGTCACTCAA 1620
Db	1561 GTGGTTTCAGTTTGTGGTGTCTGCTGATGATGATCCAGCAGTTTGCCATTGTCACTCAG 1620
QY	1621 TACATATCAGGGGTTCTCTGTTCTCCCTCTCATGAGCAGAGAGGATTTCTTCAATTG 1680
Db	1621 TACATTTCAAGAGGCTCCCTGTTCTCCCTGCTTTCATGAGCAGAGAGATTTCTTCACTG 1680
QY	1681 CAGTCTAAATTTAATTTGTCAGTATGTTGTCGAAAGCATGAGTACCTTCAACACTG 1740
Db	1681 CAGTCTAAATTTAATTTGTCGATAGTGTGTCGAAAGCATGAGTACCTTCAACACTG 1740
QY	1741 ACACAGCCAAATTTATCATCTGCTGACTTGAACAGTCAATATTTCTTCTATGAGGATGGG 1800
Db	1741 ACCAGCCAAATCATACACGCGACCTGAAAGCCACATATTTCTGCTATGAGATGGC 1800
QY	1801 CATGCTGTGGTGGCAGATTTTGGAGAAATCAAGATTTCTACAGTCTCTGGATGAAGACAAC 1860
Db	1801 CATGCTGTGGTGGCAGATTTTGGAGAAATCAAGATTTCTGCACTCCCTGGATGAAGACAAC 1860
QY	1861 ATGACAAAACACCTGGGAACTCCGTTGGATGCTCCTGAGGTTGTTCAAGCAGTGCAC 1920
Db	1861 ATGACAAAGCAGCCAGGAACTCCGCTGGATGGCCCTGAGGTTGTTCAACAGTGCACG 1920

QY	1921	CGGTACACCATCAAAAGCAGATGCTCTTACGCTATGCTCTGTCTGTGGGAAATCTCACT	1980
Db	1921	AGATACACCATCAAGGCTGATGCTTTCAGTTACTCCCTGTCTGTGGAGCTCCTCACT	1980
QY	1981	GGCGAAATTCATTTGGCTCATCTCAAGCAGCGGCTGGCGAGCAGACATGGCTTACCAC	2040
Db	1981	GGAGAAATTCATTCGCTCATCTCAAGCAGCGCTGGCGAGCAGATATGGCGTATCAC	2040
QY	2041	CACATCAGACCTCCATTTGGCTATTCATTCATTCAGAGCCCATCATCTCTCTGCTATACGA	2100
Db	2041	CACATCAGACCCCATCGGCTATTCATTCATTCAGAGCCCATCTCATCCCTCTGATACGG	2100
QY	2101	GGGTGAACGCTATGCTCTGAAGGAGACCCGAAATTTCTGAAGTTGTATGAAGTTAGAA	2160
Db	2101	GGCTGAAATGATGCTCTGAAGGAGACGACAGAGTTCTCTGAAGTTGTATGAAGTTAGAA	2160
QY	2161	GAGTGTCTCTGCAACATTTAGCTGATGCTCTGCTCATCAAGTAACAGCAGTGGGTCTCTC	2220
Db	2161	GAGTGTCTATGCAATTTAGCTGATGCTCTGCTCATCAAGTAACAGCAGTGGGTCTCTC	2220
QY	2221	TCACCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	2280
Db	2221	TCACCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	2280
QY	2281	GCAGCAATTAAGAGTGGTTTCCGAATTTGGAATATGCTCTAAATGCAAGTCTCTATGCTCT	2340
Db	2281	GCAGCCTTACGAGCGGTTTTCAGTTGGAGTATGCCCTAAATGCAAGTCTCTATGCTCTGG	2340
QY	2341	TTGTCCCAAGTCTGGACAAATTCCTCTCAAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	2400
Db	2341	TGGTCCCAAGTCTGGACAAATTCCTCTCAAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	2400
QY	2401	AGTCTTCAATACACACCCATTCACAAATATGCTATGATCCGATCCCATGAGCTCAATG	2460
Db	2401	AGCACCCAGTATTAACCTGTTGACAAATACGCTATGCTGATCCCATGAGCTGAGC	2460
QY	2461	CATTTTCATCTTGGCGAAATATGATGACGCTTTGAGGACAGCAGC	2505
Db	2461	CACCTTCACTCCCGCAAGACGACGACACTTTGAGGACAGCAAC	2505

RESULT 12  
AR437548  
LOCUS AR437548 2505 bp DNA linear PAT 18-DEC-2003  
DEFINITION Sequence 9 from patent US 6660490.  
ACCESSION AR437548  
VERSION AR437548.1 GI:40202643  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 2505)  
AUTHORS Raju, J.  
TITLE CARK protein and nucleic acid molecules and uses therefor  
JOURNAL Patent: US 6660490-A 9 09-DEC-2003;  
FEATURES  
Location/Qualifiers  
1..2505  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN  
Query Match 75.5%; Score 1892.2; DB 6; Length 2505;  
Best Local Similarity 84.7%; Pred. No. 0;  
Matches 2122; Conservative 0; Mismatches 383; Indels 0; Gaps 0;  
QY 1 ATGGGAAATTAATAATCTAGACCAACCCAACTTGTACTGATGAATGGAGAAAAATC 60  
Db 1 ATGGGAAATTAATAATCTAGACCAACCCAACTTGTACTGATGAATGGAGAAAAATC 60  
QY 61 AGTGAAATCATATTTATCAATAGAAAGATGATGACCTGACATCAGGAAAAA 120  
Db 61 AGTGAAATCATATTTATCAATAGAAAGGCTGGAGGATAAAGCTGACATCAAGAAAA 120

```

1201 CTGAAGCACTACAGAGACCCAGGAGGAGTGCATGTAAACGAATATTTCCAGCCTGGA 1260
1261 GGAGATGCTCCTATGTCTGTCTTCCATCACCCTTGGGAGAGATTAAAGCATGACAAA 1320
1261 GGAGATGCTCCTATGTCTGTCTTCCATCACCCTTGGGAGAGATTAAAGCATGACAAA 1320
1321 GAGAAGGCAGATATTTCTCCTCTAAGAGCTGGATTGCTTCCATATTTCCATCTTCAGCTC 1380
1321 GAGAAGGCAGATATTTCTCCTCTAAGAGCTGGATTGCTTCCATATTTCCATCTTCAGCTC 1380
1381 TCAGAAATGAGTTCCATGAGATATTTGGCTCAGGTTCTTTGGGAAAGATATATAAGA 1440
1381 TCCGAAATCGAGTTCCAGAGATATTCGGCTCGGGTTCTTTGGGAAAGTCTATAAGGG 1440
1441 CGATCGAGAAATAAATAGTGGCTATAAAGCTTATCGAGCCAATACCTACTGCTCCAAG 1500
1441 CGATCGAGAAATAAATAGTGGCTATAAAGCTTATCGAGCCAATACCTACTGCTCCAAG 1500
1501 TCAGATGAGATATTTTTCGCGAGAGGTGTCCATTTCTGCGAGCTCAATATCCCTGTC 1560
1501 TCAGATGAGATATTTTTCGCGAGAGGTGTCCATTTCTGCGAGCTCAATATCCCTGTC 1560
1561 GTAATTCAATTTGGGGTGTCTTGAATGATCCAGCCAGTTTCCCATTTGCTCACTAA 1620
1561 GTGGTTCAATTTGGGGTGTCTTGAATGATCCAGCCAGTTTCCCATTTGCTCACTAA 1620
1621 TACATATCAGGGGTTCTGTGTTCTCCCTCTTTCATGAGCAGAGAGGATTTCTGATTTG 1680
1621 TACATTTCAGGAGGCTCCCTGTTCTCCCTGTTTCATGAACAGAGAGATTTCTGACTTG 1680
1681 CAGTCTAAATTAATTTAGTATGATGTTGCGAAGGATGAGTACCTTCAACACCTG 1740
1681 CAGTCTAAATTAATTTAGTATGATGTTGCGAAGGATGAGTACCTTCAACACCTG 1740
1741 ACACAGCAATTTATACATCTGCTGAAACAGTCAACAATTTCTTCTATGAGGATGG 1800
1741 ACCAGCAATCATACACCGGACCTGAAACAGCACAATATTTCTGCTATGAGGATGG 1800
1801 CATGCTGTGGGAGATTTTGGAGATCAAGATTTCTACAGTCTCTGGATGAAGACAC 1860
1801 CATGCTGTGGGAGATTTTGGAGATCAAGATTTCTGAGTCCCTGGATGAAGACAC 1860
1861 ATGACAAAACACCTCGGAACCTCCGTTGATGGCTCCTGAGTGTTCACGAGTGCAT 1920
1861 ATGACAAAGCAGCCAGGAACTCCGCTGATGGCTGATGGCTGATGGCTGATGGCTG 1920
1921 CGGTACACCATCAAGCAGATGTTTGAAGTATGCTATGCTGTGTGTGGGAAATTTCTACT 1980
1921 AGATACACCATCAAGGCTGATGTTTCAAGTACTCTGCTGTGTGTGGGAGCTCTCTACT 1980
1981 GGCAGAAATTCATTCGCTCATCTCAAGCAGGCGCTGCGGACGACAGACATGCTTACCAC 2040
1981 GGCAGAAATTCATTCGCTCATCTCAAGCAGGCGCTGCGGACGACAGATGCGGTATCAC 2040
2041 CACATCAGACCTCCCATTTGCTATTCATTTCCCAAGCCCATATCATCTCTGCTGATAGA 2100
2041 CACATCAGACCTCCCATTTGCTATTCATTTCCCAAGCCCATATCATCTCTGCTGATAGA 2100
2101 GGGTGGAAACGATGCTCTGAGGAAGACCGGATTTTCTGAAGTGTCTAGAGTTAGAA 2160
2101 GGGTGGAAATGATGCTCTGAGGAAGACCGGATTTTCTGAAGTGTCTAGAGTTAGAA 2160
2161 GAGTGTCTCTGCAACATTTAGCTCATGTCTCTGATCAAGTAAACAGAGTGGGTCTCTC 2220
2161 GAGTGTCTCTGCAACATTTAGCTCATGTCTCTGATCAAGTAAACAGAGTGGGTCTCTC 2220
2221 TCACCTTCTTCTTCTGATGCTGCTGAGTGAACCGGGAGGACCTGGCGGAGTCAATGT 2280
2221 TCACCTTCTTCTTCTGATGCTGCTGAGTGAACCGGGAGGACCTGGCGGAGTCAATGT 2280
2281 GCAGCTTAAGAAGTCTGTTTCAATTTGGAATGCTCTAAATGCAAGGTCCTATGCTGCT 2340

```

2281 GCAGCTTACGAGCCGCTTTTGAGTTGGAGTATGCCTAAATGCAAGGTCCTATGCTGGG 2340

2341 TTGTCCCAAGTCTGAGCAATATTCCTCTCAAGGTCCTGCTTTTGGAGGATGAAAAA 2400

2341 TGTGCCAAAGTGTGGACACACTCTAATCGGGCTGTCTTTGGAGGATGAAATAGG 2400

2401 AGTCTTCAATACACACCCATTGACAAATATGCTATGATCCGATCCCATGAGCTCAATG 2460

2401 AGCACCCAGTATTCAACTGTTGACAAATACGCTATGCTGATCCCATGAGCTGACG 2460

2461 CATTTTCATCTTCCGAAATAGTAGCAGCTTTGAGGACAGCAGC 2505

2461 CACTTCACTCCCGCAACACGACCAACTTTGAGGACAGCAAC 2505

RESULT 13

AY303692 1982 bp mRNA linear ROD 23-JUN-2003

LOCUS AY303692

DEFINITION Rattus norvegicus cardiac ankyrin repeat kinase mRNA, complete cds.

ACCESSION AY303692

VERSION AY303692.1 GI:32165609

KEYWORDS

SOURCE Rattus norvegicus (Norway rat)

ORGANISM Rattus norvegicus

REFERENCE 1 (bases 1 to 2982)

AUTHORS Jeyaseelan,R.

TITLE Rat cardiac ankyrin repeat kinase (rCARK)

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 2982)

AUTHORS Jeyaseelan,R.

TITLE Direct Submission

JOURNAL Submitted (20-MAY-2003) Cardiovascular Biology, Millennium Pharmaceuticals Inc., 75 Sidney Street, Cambridge, MA 02139, USA

FEATURES

Location/Qualifiers

1..2982

/organism="Rattus norvegicus"

/mol\_type="mRNA"

/db\_xref="taxon:10116"

44..2551

/note="CARK; protein kinase"

/codon\_start=1

/product="cardiac ankyrin repeat kinase"

/protein\_id="AAP72031.1"

/db\_xref="GI:32165610"

/translation="MGNYKSRPTQCTDEWKKKVSEYAIIELEDNLQIKENEFQELRHIFGSDAESEVSLNRYTERGSLHLLCCCGNKKSHRALMLKGLRPSRLTRNGFPAHLAVYKDSPELITSLHSGADVQVQVGLTALHIAIAGHPEAAEVLQHGAVNVQDAVFTPLHIAAYVGEQVTVLLKFGADVNSVGEVDRPLHLASAKGFFNVKLIVYEGSKADVNAQDNEDHVLPHFGSRFGHHNI VSYLLQSDLEVPVHLYGDPDLHLACYNQFVAKEIVQVGTESLITENIFSETAFHSACTYGNLIDLVFLDQNAVINHRGDHGTGLHSACTGHIRLVFLDNDGADNVLVADPSSRSGEKDEQCTCLMVALEKGDHAI VTLTKHYRQPEELPCNYSQPGDGSVSVSPGLKIKSMTKKADVLLURAELPSRFHQLQSEIEFHEIIGSGFKYKGRNKVIAIKRYANTYCSKSDVMFOR EYSLICQLNHPVQVQVGLDPSQFAIVTQYISGSLFSLHQRRLIDLSKLII AVDAVKMEYHLSTLQPIIHRDLNSHNILLYEDGHAVVADFGESRFLQSLDENMTKQPCNLAWAPVPTOCTRYTIKADVFSYSLCLWELLTGEIIPAHLKPAAAAADVAHHIRPPIGYSIPKPISSLLIRGNWACPEGRPEFSEVSVKLEECCLNVELMSPADSNSSGSLSPSSSDCLLRGGGPRSHVAALRSFELETALNARSYAGSQSVGTHSPFGLSLEEMNRSTQYSTVDKYGVYVSDPMSLTHLSHQDDNSDFDSN"

Query Match 75.5%; Score 1892.2; DB 10; Length 2982;

Best Local Similarity 84.7%; Pred No. 0;

Matches 2122; Conservative 0; Mismatches 383; Indels 0; Gaps 0;

QY 1 ATGGGAAATTAATAATCTAGACCAACCCAACTTGCTGATGATGAGAGAAAAAGTC 60

Db 44 ATGGGAAATTAATAATCCAGCAACACAGACTTGTCTGATGATGAGAGAAAGATT 103

QY 61 AGTGAATCATATGTTTATCACAATAGAAAGATTAGAGATGACCTGCAGATCAAGAAAA 120

Db 104 AGTGAATCTTACGCTATTTATCATAGAAAGGCTGGAGGATACCTGCGAGATCAAGAAAT 163  
Qy 121 GAACCTGACAGAACTAAGGAATATATTTGGCTGTGATGAAGCCTTCAGTAAGTCAATTTA 180  
Db 164 GAATTTCAAGAACTAAGGCACATCTTTGGCTCTGATGAAGCCTTCAGTGAAGTCAAGTTA 223  
Qy 181 AATTACCGCACTGAAAATGGGCTGTCTCTACTTCAATTTATGTTGATTTGAGGAGCAAG 240  
Db 224 AATTACCGCACAGAGGCTGGCCCTGTCTGTCTACACCTCTGCTGTCTGTGGGGCAAC 283  
Qy 241 AAATCAATATTCGAACCTCTTATGTTGAAGGGCTCCGCCCATCTCGACTGCAAGAAAT 300  
Db 284 AAGTCACATATCCGTGCTCTTATGTTAAAGGGCTCCGTCATCCAGACTGACGAGAAAT 343  
Qy 301 GGNATTACAGCCTTGCAATTTAGACGTTTACAGAGGATATGAGAAATGATCAGCTTCTCTG 360  
Db 344 GGGTTTCCAGCTCTGCACTGCGCCGTTTACAGGACGCGCGGAACCTTATCAGTTCACTG 403  
Qy 361 CTTACAGTGGAGCTGATATACAGCAGGTTGGATACGGTGGCCCTCACTGCCCTCCATATT 420  
Db 404 TTGCACAGCGAGCAGATGTTTCAGCAAGTGGGATACGGTGGCCCTCACAGCCCTCCACATA 463  
Qy 421 GCTACAAATAGCTGGCCACCTAGAGGCTGCTGATGTGCTGTTGGCAACATGAGCTAAATGTC 480  
Db 464 GCTGCAATAGCTGGACACCCAGAGGCTGCAAGAGTGTGCTACAAACATGGGGCCAAATGTG 523  
Qy 481 AATATTCAAGATGTCAGTGTCTTTTCACTCCATTTGATATTTGAGCGCTACTATGGACATGAA 540  
Db 524 AATGTTCAAGATGCGGCTTCTTCAACCCACATGACATTCAGGCTACTATGGGCACGAG 583  
Qy 541 CAGGTAACCTCGCCTTTCTTTTGAATTTGGTGTGATGTAATGTAAGTGGTGAAGTTGGA 600  
Db 584 CAGGTAACCACTGTCCTTTTGAAGTTGGTGTGATGTCAATGTAAGCGGTGAAGTTGGG 643  
Qy 601 GATAGACCCCTCCACCTAGACATCTGCAAGGATTTGAAATATGTAATGTAAGGATCTTGTATG 660  
Db 644 GACAGGCTCTGCACTGCGCCTCTGCAAGGCTTCTTCAACATTTGAAATCTCTGCTGTA 703  
Qy 661 GAAGAGGCGAGCAAAAGCAGATGTAATGCTCAAGATAATGAAGACCATGTCCCACTCCAT 720  
Db 704 GAAGAGGCGAGCAAAAGCAGATGTAAGCTCAGGACCAATGAAGACCATGCTCCCTCTGCAC 763  
Qy 721 TTCTGTTCTGATTTGGACACCATGATAGTAAATGTAAGTGAAGTGAAGTGAAGTGAAG 780  
Db 764 TTCTGTTCTGATTTGGACACCAATATAGTGAAGTACCTGCTCCAGAGTGAATGAG 823  
Qy 781 GTTCAACCTCATGTTGTTAATATCTATGGAGATACCCCTTACACCTGGCATGCTACAAT 840  
Db 824 GTCCAGCCTCAGCTCAATTAACATCTATGTTGACATCTCTTTGCACCTGGCATGCTACAAT 883  
Qy 841 GGCAAAATTTGAAGTGGCAAGGAATCATCCAAATATCAGGAAACAGAAAGTCTGACTAAG 900  
Db 884 GGAAATTTTGAAGTGGCAAGGAATTTGTCAGGTAAACAGGAACCTGAAAGTCTGACTAAG 943  
Qy 901 GAAACATCTTCAGTGAACACGCTTTTCATAGTGTGTTGACCTATGGCAAGCAGATTGAC 960  
Db 944 GAAACATCTTCAGCAGAGACGCTTTTCAGATGCTTTGACCTATGGCAGAAACATTTGAC 1003  
Qy 961 CTAGTCAAAATTTCTTCTGATCAGATGTCATAAATCAACACCAAGGAAGGATGGG 1020  
Db 1004 CTGCTCAAAATTTCTTCTGATCAGATGCTGTGAACATTAACCAACCGAGGAAGATGGG 1063  
Qy 1021 CACACTGGATATACATCTGCTGCTACAGGCTCAGATTCGCTGGTTCAGTTCTTACTG 1080  
Db 1064 CACACAGGATTGCACTCTGCTTCTACACGCGCATATCCGCTGCTGCTGCTGCTTCTTCT 1123  
Qy 1081 GATAATGGAGCTGATATGAATCTAGTGGCTTGTGATCCAGCAGGCTCTAGTGGTGAAGAA 1140  
Db 1124 GATAATGGTGCATATGAATCTTGTGCTGTTGATCCAGCAGGCTCTAGTGGTGAAGAA 1183  
Qy 1141 GATGACAGACATGTTGTGATGTTGGCTTATGAAAGGGGCGATGATGCCATTTGTCACATC 1200

Db 1184 GATGACGACAGCATGTTTGTATGGGCTTTACAGAAAGGACATGATGCCATTTGTTACACTC 1243  
Qy 1201 CTGAAGCATTTAAGAGACCAAGATGTAATGGCTTGTATGAATATTTCTCAGCCCTGGA 1260  
Db 1244 CTGAAGCATTTAAGAGACCAAGAGGAGCTGCATGTAACGATATTTCCAGCCTGGA 1303  
Qy 1261 GGAGATGCTCTCTATGTTCTGTGTTCCATCACCTTTGGGGAAGATTTAAAGCATGACAAA 1320  
Db 1304 GGAGATGCTCTCTATGTTCTGTTCCTTCCCTTTGGGCAAGATTTAAAGCATGACAAA 1363  
Qy 1321 GAGAGGCGAGATATCTCTCTAAGAGCTGGATGCTTCCACATTTCCATTTCTTCACTC 1380  
Db 1364 GAGAGGCGAGATGTTCTCTCTGAGGCTGAATCACTACCTCCCTTCCATCTTCCAACTC 1423  
Qy 1381 TCAGAAATTTGAGTTCCATGAGATTTTGGCTCAGGTTCTTTTGGGAAAGTATATAAGGA 1440  
Db 1424 TCCGAAATCGAGTTCCACAGATTTATGGCTCGGTTCTTTTGGGAAAGTCTATAAGGG 1483  
Qy 1441 CGATCGAGAAATAAATAGTGTCTATAAAAGTTATCGAGCCATAGCTACTGCTCCAG 1500  
Db 1484 CGATCGAGAAATAAATAGTGTGTGCGAGAGTGTCCATCTCTGCCAGCTCAACCCCTGC 1543  
Qy 1501 TCAGATGTGATATGTTTGGCGAGAGTGTCCATTTCTCTCCAGCTCAATCATCCCTGC 1560  
Db 1544 TCAGAGCTGATATGTTTGGCGAGAGTGTCCATCTCTGCCAGCTCAACCCCTGC 1603  
Qy 1561 GTAAATTCAGTTTGGTGTGTTGATGATCCAGCAGTTTGCATTTGCTTCACTCAA 1620  
Db 1604 GTGGTTTCACTTTGGTGTGCTGCTGATGACCCAGTCACTTTGCTTGTCACTCAG 1663  
Qy 1621 TACATATCAGGGGTTCTCTGTTCTCCCTCTCATGAGCAGAGAGATTTCTTGAATTTG 1680  
Db 1664 TACATTTCAAGAGGCTCCCTGTTTCTCCCTGTTTATGAAACAGAGAGATTTCTTGAATTTG 1723  
Qy 1681 CAGTCTTAAATTAATTTGCACTAGATTTGCCAAAGGCATGGAGTACCTTTCAACACTG 1740  
Db 1724 CAGTCTTAAATTAATTTGCGGTAGACGTTGCCAAGGCGATGGAGTACCTGCACAGTTG 1783  
Qy 1741 ACAGCCCAATTAACATGCTGATGTAACAGTCAAAATTTCTTCTATGAGAGTGGG 1800  
Db 1784 ACCAGCCCAATTAACATGCTGATGTAACAGGCTGTAACAGGCTTCTTATGAGAGTGGC 1843  
Qy 1801 CATGCTGTGGTGGCAGATTTTGGAGATCAAGATTTCTGAGTCTCTGAGTGAAGACAAAC 1860  
Db 1844 CATGCTGTGGTGGCAGATTTTGGAGATCAAGATTTCTGAGTCTCTGAGTGAAGACAAAC 1903  
Qy 1861 ATGACAAAACAACTGGGAAACCTCGTTGGATGGCTCTGAGGTTTCCAGCAGTGCAC 1920  
Db 1904 ATGACAAAACAACTGGGAAACCTCGCTGGATGGCTGGATGGCTGGATGGCTGGATGG 1963  
Qy 1921 CGGTACACCATCAAGCAGATGCTTTCACTGATGCTCTGCTGTGGGAAATTTCTCACT 1980  
Db 1964 AGATCACCATCAAGCTGATGTTCTCAGTTACTCTCTGTTCTGTGGAGCTCTCACT 2023  
Qy 1981 GGCGAAATTTCCATTTCCCTCATCTCAAGCCAGCGGTGGGCGAGCAGATGGCTTACAC 2040  
Db 2024 GGAGAAATTTCCATTTCCCTCATCTCAAGCCAGCGGTGGGCGAGCAGATGGCTTACAC 2083  
Qy 2041 CACATCAGACCTCCCATTTGGCTATTTCCATTTCCAGCCCATATCATCTCTGATAGGA 2100  
Db 2084 CACATCAGACCGCCCATTTGGCTATTTCCATTTCCAGCCCATCTCATCTCTGATAGCG 2143  
Qy 2101 GGGTGAACGCGATGCTTGAAGGAAGACCCGAATTTCTGAAGTTGTCTGAAGTTAGAA 2160  
Db 2144 GGCTGAATGCTATGCTTGAAGGACGACAGATTTCTCTGAAGTCTGTAGCAAACTGGAG 2203  
Qy 2161 GAGTGTCTCTGCAACATTTAGAGCTGATGCTCTCTGATCAAGTAACAGCAGTGGGTCTCTC 2220  
Db 2204 GAGTGTCTTATGCAATTTGAGGCTCATGCTCCAGCATCAAGTAACAGCAGTGGGTCTCTG 2263  
Qy 2221 TCACCTTCTTCTTCTTCTGATTTGCTGTAACCGGAGGACCTGGCCGAGGATCATGTG 2280  
Db 2264 TCACCTTCTTCTTCTTCTGATTTGCTGAGCCCGGAGGAGGCTGGCCGAGGACCATGTG 2323



Qy 1261 GGAGATGGCTCCTATGTGTGTTCCATCAACCCCTGGGGAAGATTAAAGCATGACAAAA 1320  
Db 1321 GGAGATGGCTCCTATGTGTGTTCCATCAACCCCTGGGGAAGATTAAAGCATGACAAAA 1380  
Qy 1321 GAGAAGCAGATATTCCTCTCTTAAGAGCTGATTCCTCTTCAATTTCCATCTTCAGTTC 1380  
Db 1381 GAGAAGCAGATATTCCTCTCTTAAGAGCTGATTCCTCTTCAATTTCCATCTTCAGTTC 1440  
Qy 1381 TCAGAAATTCAGATTCATGAGATTAATGGCTCAGGTTCTTTTGGGAAAGATATATAAGGA 1440  
Db 1441 TCCGAATTCAGATTCATGAGATTAATGGCTCAGGTTCTTTTGGGAAAGATATATAAGGA 1500  
Qy 1441 CGATGCAGAAATAAATAGTGGCTATAAAGCTTATCGAGCAATACCTACTGCTCCAAAG 1500  
Db 1501 CGATGCAGAAATAAATAGTGGCTATAAAGCTTATCGAGCAATACCTACTGCTCCAAAG 1560  
Qy 1501 TCAGATGTGATATGTTTCCGAGAGGTGTCATTTCTCTCCCTCTCATGACAGAGGATTTCTTGAATG 1560  
Db 1561 TCAGATGTGATATGTTTCCGAGAGGTGTCATTTCTCTCCCTCTCATGACAGAGGATTTCTTGAATG 1620  
Qy 1561 GTAAATTCAGTTCGTGGGTGCTTCTTGAATGATCCAGCCAGTTTGCCATTTGTCACTCAA 1620  
Db 1621 GTGTTTCAGTTCGTGGGTGCTTCTTGAATGATCCAGCCAGTTTGCCATTTGTCACTCAA 1680  
Qy 1621 TACATATCAGGGGTTCTCTGTTCTCCCTCTCATGACAGAGGATTTCTTGAATG 1680  
Db 1681 TACATATCAGGGGTTCTCTGTTCTCCCTCTCATGACAGAGGATTTCTTGAATG 1740  
Qy 1681 CAGTCTAAATTAATTCAGATAGATGTTGCCAAAGGATGAGTACCTTCACAACTG 1740  
Db 1741 CAGTCTAAATTAATTCAGATAGATGTTGCCAAAGGATGAGTACCTTCACAACTG 1800  
Qy 1741 ACACAGCAATATACATCGTGAATTTGAACAGTCAATATTTCTTCTATGAGGATGG 1800  
Db 1801 ACCACAGCAATATACATCGTGAATTTGAACAGTCAATATTTCTTCTATGAGGATGG 1860  
Qy 1801 CATGCTGTGTGCAGATTTTGGAGATCAAGATTTCTACAGTCTCTGATGAAGACAAAC 1860  
Db 1861 CATGCTGTGTGCAGATTTTGGAGATCAAGATTTCTACAGTCTCTGATGAAGACAAAC 1920  
Qy 1861 ATGCAAAACCAACCTGGAAACCTCCGTTGATGGCTCTTCTGAGTGTTCACGAGTGCAT 1920  
Db 1921 ATGCAAAACCAACCTGGAAACCTCCGTTGATGGCTCTTCTGAGTGTTCACGAGTGCAC 1980  
Qy 1921 CGGTACACCATCAAGCAGATGTTCTCAGTATGCTCTGCTGTGTGGAAATTTCTCACT 1980  
Db 1981 AGATACACCATCAAGCAGATGTTCTCAGTATGCTCTGCTGTGTGGAAATTTCTCACT 2040  
Qy 1981 GGCAGAAATTCATTCGCTCATCTCAAGCCAGCGCTGCGGACGACACATGCTTACCAC 2040  
Db 2041 GGAGAAATTCATTCGCTCATCTCAAGCCAGCGCTGCGGACGACATGCTTACCAC 2100  
Qy 2041 CATATCAGACTCCCATGCTATTCATTCACAGCCCATATCATCTCTGCTGATACGA 2100  
Db 2101 CATATCAGACTCCCATGCTATTCATTCACAGCCCATATCATCTCTGCTGATACGA 2160  
Qy 2101 GGGTGGAAACCATGCTCTGAGGAGACACCGAATTTCTGAAATGCTCATGAAGTTAGAA 2160  
Db 2161 GGGTGGAAACCATGCTCTGAGGAGACACCGAATTTCTGAAATGCTCATGAAGTTAGAA 2220  
Qy 2161 GAGTGTCTCTGCAACATGAGCTGATGTTCTCTGATCAAGTAAACAGCAGTGGGTTCTC 2220  
Db 2221 GAGTGTCTCTGCAACATGAGCTGATGTTCTCTGATCAAGTAAACAGCAGTGGGTTCTC 2280  
Qy 2281 TCACCTTCTCTCTGATTCGCTGGTGAACCGGGGAGCCTGGCGGAGTCAATG 2280  
Db 2281 TCACCTTCTCTCTGATTCGCTGGTGAACCGGGGAGCCTGGCGGAGTCAATG 2340  
Qy 2281 GCAGCATTAAGAGTCCGTTTGAATGGAATATGCTTAAATGCAAGGTCCTATGCTGT 2340  
Db 2341 GCAGCATTAAGAGTCCGTTTGAATGGAATATGCTTAAATGCAAGGTCCTATGCTGT 2400

Qy 2341 TTGTCCTCAAGTGTCTGGCAATATTCCTCTCAAGTGTCTGTTTGGAGGAGATGAAAGA 2400  
Db 2401 TGTGTCCTCAAGTGTCTGGCAACACATCTAATCCGGGCTGTCTTTGGAGGAGATGAAATAG 2460  
Qy 2401 AGTCTTCAATACACACCCATTCACAAATATGCTATGATCCGATCCCATGAGCTCAATG 2460  
Db 2461 AGCACCAGATTTCAATGTTGACAAATACGGCTAATGCTGATCCCATGAGCCTGAG 2520  
Qy 2461 CATTTTCATTTCTGCGAAATATGATGACGCTTTGAGGACAGCAGC 2505  
Db 2521 CACCTTCACTCCGCGCAACAGCAGCAACTTTGAGGACAGCAAC 2565

RESULT 15  
AR269113 LOCUS 3026 bp DNA linear PAT 10-APR-2003  
DEFINITION Sequence 7 from patent US 650654.  
ACCESSION AR269113  
VERSION AR269113.1 GI:29699956  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 3026)  
AUTHORS Raju, J.  
TITLE CARX protein and nucleic acid molecules and uses therefor  
JOURNAL Patent: US 650654-A 7 31-DEC-2002;  
FEATURES Location/Qualifiers  
source 1..3026  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN  
Query Match 75.5%; Score 1892.2; DB 6; Length 3026;  
Best Local Similarity 84.7%; Pred. No. 0;  
Matches 2122; Conservative 0; Mismatches 383; Indels 0; Gaps 0;

Qy 1 ATGGGAATATATAAATCTAGACCAACCCCAAACTGTACTGATGAATGGAAGAAAAAGTC 60  
Db 61 ATGGGANATTACAAATCCAGACCAACACAGACTTGTCTGATGAATGGAGAAGAGTT 120  
Qy 61 AGTGAATCATATGTTATCAATAGAAAGATTAGAAGATGACCTGCAGATCAAGGAAAA 120  
Db 121 AGTGAATCTTTACGCTATTATCATAGAAGGCTGGAGGATAACCTGCAGATCAAGAAAA 180  
Qy 121 GAATGACAGAACTAAGGAATATATTGGCTCTGATGAAGCCTTCAGTAAAGTCAATTTA 180  
Db 181 GAATTTCAAGAACTAAGGACATCTTTGGCTCTGATGAAGCCTTCAGTGAAGTCAATT 240  
Qy 181 AATTACCGCACTGAAATGGGCTGCTCTACTTCAATTTATGTTGCAATTTGTGGAGCAAG 240  
Db 241 AATTACCGCACAGCGTGGCCTGCTCCTGTACACCTCTGCTGTCTGTGGCGCAAC 300  
Qy 241 AATCATCATATTGCAACTCTTATGTTGAAGGCTCCGCCATCTCGACTGACAGAAAT 300  
Db 301 AAGTCATATATCCGTGCCCTTATGTTTAAAGGGCTCCGTCCTCAGACTGACAGAAAT 360  
Qy 301 GGATTTACAGCCTTGCAATTTAGCAGTTTACAAGGATAATGACAGAAATTTGATCACTTCT 360  
Db 361 GGGTTTCCAGCTCTGACCTGGCCGTTTACAAGGACAGCCCGGAACCTTATCACTTCACT 420  
Qy 361 CTTACAGTGGAGCTGATATACAGCAGGTTGATACGGTGGCCTCACTGCCCTCCATATT 420  
Db 421 TTGCACAGCGAGCAGATGTTCAGCAAGTGGGATACGGTGGCCTCACAGCCCTCCACATA 480  
Qy 421 GCTACAAATAGCTGGCACCTAGAGGCTGCTGATGCTGTTGCTTCAACATGGAGCTAATGTC 480  
Db 481 GCTGCAATAGCTGGACACCCAGAGGCTGCAGAAGTGTGCTACACATGGGGCCATGTG 540  
Qy 481 AATATTCAAGATGCAATTTTCTTCACTCCATTCATTTGACAGCTACTATGGAATGAA 540  
Db 541 AATGTTCAAGATGCCGCTCTTCTTCACTCCACTGCACATTTGACGCTACTATGGGACGAG 600

QY 541 CAGGTAACCTCGCCTTCCTTTGAAATTTGGTCTGATGTAAATGTAAGTGTGAAGTTGGA 600  
DB 601 CAGGTAACCTCGCCTTCCTTTGAAATTTGGTCTGATGTAAATGTAAGTGTGAAGTTGGA 660  
QY 601 GATAGACCCCTCCACTAGCATCTGCAAAAGGATTTCTTGAATATTGCAAACTCTTTGATG 660  
DB 661 GACAGCCCTCTGCACCTCGCCTCTGCAAGGCTTCTTCAACATTGTGAACCTCTCTGTA 720  
QY 661 GAAGAGGAGCAAGCAGATGTGAATGCTCAAGATTAATGAAGCAGATGCCCACCTCCAT 720  
DB 721 GAAGAGGAGCAAGCAGATGTGAAGCTCAGACCAATGAAGACCAAGCTCCCTCTGCAC 780  
QY 721 TTCTGTTCTCGATTTGGACACCATGATATAGTTAAGTATCTGTGCAAGAGTATTTGGAA 780  
DB 781 TTCTGTTCTCGATTTGGACACCAATATAGTTAGCTACTGCTCCAGAGTACTTAGAG 840  
QY 781 GTTCAACCTCATGTTGTTAATCTATGGAGATACCCCTTTACACCTGGCATCTACAAT 840  
DB 841 GTCCAGCCTCAGTCAATTAACATCTATGTTGACACTCTTTTGACCTGGCATCTACAAT 900  
QY 841 GCAAAATTTGAAGTTGCCAAGAAATCATCCAAATATCAGAAACAGAAATCTGACTTAG 900  
DB 901 GCAAAATTTGAAGTTGCCAAGAAATTTGTCAGGTAAACAGAAATGAAAGTCTGACTAAG 960  
QY 901 GAAACATCTTCAGTGAAACAGCTTTTCATAGTCTTTGTACTATGGCAAGAGCATTTGAC 960  
DB 961 GAAACATCTTCAGTGAAACAGCTTTTCACAGTCTTTGTACTATGGCAAGAAATTTGAC 1020  
QY 961 CTAGTCAAAATTTCTTTGATCAGAAATGTCAAAATCAAAATCAACCAAGAGGATGGG 1020  
DB 1021 CTGGTCAAAATTTCTTTGATCAGAAATGTGTGAACATTAACCAAGAGGATGGG 1080  
QY 1021 CACACTGATTTACACTCTGCTTGTCCACAGGTCAATTCGCTGTTTCAAGTCTTACTG 1080  
DB 1081 CACACAGATTTGCACTCTGCTTGTCTACAGGCAATATCCGCTGTTTCAAGTCTTACT 1140  
QY 1081 GATAATGAGTGTATGAATCTAGTGGCTTGTATCCAGCAGGCTAGTGGTGAATAA 1140  
DB 1141 GATAATGAGTGTATGAATCTAGTGGCTTGTATCCAGCAGGCTAGTGGTGAATAA 1200  
QY 1141 GATGAGCAGATTTTGTGCTGCTTATGAAGAGGATGATGCCATTTGACACTC 1200  
DB 1201 GATGAGCAGATTTTGTGCTGCTTATGAAGAGGATGATGCCATTTGACACTC 1260  
QY 1201 CTGAAGCATTTAAGAGACCAAGATGAATGCTGCTGTAATGAATATTTCTAGCCTGGA 1260  
DB 1261 CTGAAGCATTTAAGAGACCAAGAGGATGCTGCTGTAATGAATATTTCTAGCCTGGA 1320  
QY 1261 GGAGTGGCTCTATGTTGCTGCTTCCATCACCCTTGGGAGGATTAAGAGCATGACAAA 1320  
DB 1321 GGAGTGGCTCTATGTTGCTGCTTCCATCACCCTTGGGAGGATTAAGAGCATGACAAA 1380  
QY 1321 GAGAAGGAGATTTCTCTCTAAGAGCTGGATTTGCTTTCATATTTCCATCTTCAGCTC 1380  
DB 1381 GAGAAGGAGATTTCTCTCTAAGAGCTGGATTTGCTTTCATATTTCCATCTTCAGCTC 1440  
QY 1381 TCAGAAATTTGATTTCCATGAGATTTGCTGCTAGGTTCTTTTGGGAAAGTATATAAGGA 1440  
DB 1441 TCCGAAATTCGAGTTCCAGAGATTTATCGGCTCGGGTTCTTTTGGGAAAGTCTATAAGGG 1500  
QY 1441 CGATGCAAAATAAATAGTGGCTATAAAGCTTTATCGAGCCTAATACCTACTGCTCCAAG 1500  
DB 1501 CGATGCAAAATAAATAGTGGCTATAAAGCTTTATCGAGCCTAATACCTACTGCTCCAAG 1560  
QY 1501 TCAGATGAGGATATTTTTCGAGAGAGGTGCTCAATTTCTGCGAGCTCAATATCCCTGC 1560  
DB 1561 TCAGAGCTGGATATTTTTCGAGAGAGGTGCTCAATTTCTGCGAGCTCAATATCCCTGC 1620  
QY 1561 GTAAATTCAGTTTGGGTGCTGCTGAATGATCCCAAGCAGTTTCCCATTTGTCACCTAA 1620  
DB 1621 GTGGTTGAGTTTGGGTGCTGCTGAATGATCCCAAGCAGTTTCCCATTTGTCACCTAG 1680  
QY 1621 TACATATCAGGGGGTTCTGTTCTCCCTCTCTCATGAGCAGAAAGGATTTCTGATTTG 1680

DB 1681 TACATTTCCAGAGGCTCCCTGTTCTCCCTGCTTTCATGAACAAGAGAAATTTCTGACTTG 1740  
QY 1681 CAGTCTAAATTAATTTAGTAGATTTGCCAAAGGATGGAGTACCTTCAACAACCTG 1740  
DB 1741 CAGTCTAAATTAATTTAGTAGATTTGCCAAAGGATGGAGTACCTTCAACAACCTG 1800  
QY 1741 ACACAGCCAAATTAATTTAGTAGATTTGCCAAAGGATGGAGTACCTTCAACAACCTG 1800  
DB 1801 ACCAGCCAAATTAATTTAGTAGATTTGCCAAAGGATGGAGTACCTTCAACAACCTG 1860  
QY 1801 CATCTGCTGGTGGCAGATTTTGGAGAAATCAAGATTTTCTACAGTCTCTGGATGAACAAC 1860  
DB 1861 CATCTGCTGGTGGCAGATTTTGGAGAAATCAAGATTTTCTACAGTCTCTGGATGAACAAC 1920  
QY 1861 ATGCAAAAACAACTGGGAACCTCCCTGGATGGCTCTGAGGTGTTACGCGAGTGCAT 1920  
DB 1921 ATGCAAAAACAACTGGGAACCTCCCTGGATGGCTCTGAGGTGTTACGCGAGTGCAT 1980  
QY 1921 CGGTACACCAATCAAGCAGATGTTTCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1980  
DB 1981 AGATACCAATCAAGCAGATGTTTCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2040  
QY 1981 GGGGAAATTCATTTCCCTCATCTCAAGCAGCGGCTGGGAGCAGAGATGCTTACCAAC 2040  
DB 2041 GGGGAAATTCATTTCCCTCATCTCAAGCAGCGGCTGGGAGCAGAGATGCTTACCAAC 2100  
QY 2041 CACATCAGACCTCCCTTCCCTCATCTCAAGCAGCGGCTGGGAGCAGAGATGCTTACCAAC 2100  
DB 2101 CACATCAGACCTCCCTTCCCTCATCTCAAGCAGCGGCTGGGAGCAGAGATGCTTACCAAC 2160  
QY 2101 GGGTGAACCACTGCTCTGAAAGAGACCCGAAATTTCTGAAGTCTCATGAAGTTAGAA 2160  
DB 2161 GGGTGAACCACTGCTCTGAAAGAGACCCGAAATTTCTGAAGTCTCATGAAGTTAGAA 2220  
QY 2161 GAGTGTCTCTGAAACATTTGAGCTGATGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2220  
DB 2221 GAGTGTCTCTGAAACATTTGAGCTGATGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2280  
QY 2221 TCACCTTCTTCTTCTTCTGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2280  
DB 2281 TCACCTTCTTCTTCTTCTGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2340  
QY 2281 GCAGCATTAAGAGAGTCTGTTTGGAAATGCTCTTAAATGCAAGTCTCTATGCTGCT 2340  
DB 2341 GCAGCATTAAGAGAGTCTGTTTGGAAATGCTCTTAAATGCAAGTCTCTATGCTGCT 2400  
QY 2341 TTGTCCCAAGTCTGCAAAATTTCTCTCAAGGTTCTGTTTGGAGAGATGAAAGA 2400  
DB 2401 TTGTCCCAAGTCTGCAAAATTTCTCTCAAGGTTCTGTTTGGAGAGATGAAAGA 2460  
QY 2401 AGTCTTCAATPACACCACTTGAATAATGCTGATGATGCTGCTGCTGCTGCTGCTGCTGCT 2460  
DB 2461 AGCACCAGTATTTCAACTGTTGCAAAATACGGCTATGTTCTGATGCTGCTGCTGCTGCTGCT 2520  
QY 2461 CATTTTCATCTTGGCCGAATAGTAGCAGCTTTTGGAGCAGCAGC 2505  
DB 2521 CACCTTCACTCCCGCCCAAGACGACGAACTTTGAGGACAGCAAC 2565

Search completed: September 6, 2004, 05:03:19  
Job time : 6448.43 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 2, 2004, 14:16:37 ; Search time 46 Seconds  
(without alignments)  
1746.085 Million cell updates/sec

Title: US-10-626-173-2

Perfect score: 4390

Sequence: 1 MGNYSRPTQTDEWKKV.....PMSSMHFHSRNSSPEDSS 835

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 78:\*\*

1: pir1:\*\*

2: pir2:\*\*

3: pir3:\*\*

4: pir4:\*\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	1667.5	38.0	842	2 T32258	hypothetical prote
2	486	11.1	546	2 D84555	probable protein k
3	475	10.8	438	2 C86273	protein kinases ho
4	473.5	10.8	553	2 T04683	hypothetical prote
5	465	10.6	1584	2 T18276	protein-tyrosine k
6	460.5	10.5	1856	2 B35049	ankyrin 1, erythro
7	460.5	10.5	1880	2 A35049	ankyrin 1, erythro
8	460.5	10.5	1881	1 SJHUK	ankyrin 1, erythro
9	457.5	10.4	1862	2 T49502	ankyrin - mouse
10	450.5	10.3	410	2 B35670	protein-tyrosine k
11	444.5	10.1	390	2 T01451	protein kinase hom
12	444	10.1	1848	2 S37771	ankyrin, erythrocy
13	443	10.1	1549	2 T13940	ankyrin - fruit fl
14	439.5	10.0	462	2 S29851	protein kinase 6 (
15	438	10.0	412	2 T10671	protein kinase hom
16	435	9.9	3924	2 S37431	ankyrin 2, neurona
17	433	9.9	407	2 G84635	probable protein k
18	431.5	9.8	1765	2 T42714	ankyrin 3, splice
19	431.5	9.8	1940	2 T42715	ankyrin 3, splice
20	431.5	9.8	1943	2 T42713	ankyrin 3, splice
21	431.5	9.8	1961	2 T42716	ankyrin 3, splice
22	427.5	9.7	4377	2 A55575	ankyrin 3, long sp
23	419	9.5	567	2 JC5957	transforming growt
24	419	9.5	579	2 JC5956	transforming growt
25	419	9.5	606	2 JC5956	transforming growt
26	414	9.4	954	1 S68178	mixed-lineage prot
27	410.5	9.4	545	2 T05675	hypothetical prote
28	408.5	9.3	1338	2 T18287	protein-tyrosine k
29	406.5	9.3	2039	2 T15347	ankyrin-related un

30 404.5 9.2 821 2 T48400 serine/threonine-p  
31 401.5 9.1 1030 2 F96763 hypothetical prote  
32 400.5 9.1 529 2 T07406 probable protein k  
33 399.5 9.1 370 2 T46150 protein kinase ATN  
34 395 1248 2 B96827 hypothetical prote  
35 394.5 9.0 364 2 G71410 probable protein k  
36 393.5 9.0 356 2 S61766 protein kinase ATN  
37 392 8.9 391 2 T48115 protein kinase ATN  
38 390.5 8.9 475 2 T12955 probable protein k  
39 389 8.9 451 1 S68455 serine/threonine-s  
40 384.5 8.8 406 2 T52626 probable mitogen-a  
41 381 8.7 466 2 T33574 hypothetical prote  
42 379 8.6 1079 1 T8FVMI gag-Rml1-env polyp  
43 379 8.6 1147 2 F86297 hypothetical prote  
44 378 8.6 736 2 T05137 protein kinase hom  
45 378 8.6 1171 2 T12956 hypothetical prote

#### ALIGNMENTS

##### RESULT 1

T32258 hypothetical protein C24A1.3 - Caenorhabditis elegans

C/Species: Caenorhabditis elegans

C/Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 29-Oct-1999

C/Accession: T32258

R/Connell, M.

submitted to the EMBL Data Library, September 1997

A/Description: The sequence of C. elegans cosmid C24A1.

A/Reference number: Z21141

A/Accession: T32258

A/Status: preliminary; translated from GB/EMBL/DBDJ

A/Molecule type: DNA

A/Residues: 1-842 <CON>

A/Cross-references: EMBL:AF024491; PIDN:AB70312.1; GSPDB:GN00021; CBSP:C24A1.3

A/Experimental source: strain Bristol N2; clone C24A1

C/Genetics:

A/Gene: CBSP:C24A1.3

A/Map position: 3

A/Introns: 43/2; 189/3; 234/1; 351/1; 511/3; 578/3; 594/2; 636/2; 675/3

Query Match 38.0%; Score 1667.5; DB 2; Length 842;  
Best Local Similarity 41.4%; Pred. No. 1.4e-81;  
Matches 367; Conservative 153; Mismatches 235; Indels 131; Gaps 22;

Qy 10 OTCTDEWKKVSSVYVITIERLEDDLOIK-----EKELTELRNIFGSDEAFS 56  
Db 2 RTSDBLKKKISGYSVRSRLSDDVRSNLGWVDVQIAAEKSLDFK-----QHMC 55  
Qy 57 KVNLYRTENGLSLHLCCICGG-----KSHIRTLMLKGLRPSRLTR 99  
Db 56 PENAELKSTQLISLPHI--ICAGHSQPEKLQFLIDLNPKESSI-TLI-----SSQSA 106  
Qy 100 NGFTALHLYKDNALITSLHSG-ADIQQYGVGLTALHIATAGHLEAADVLLOHGA 158  
Db 107 NGFTPLHLYRGDVALKATKLVLDQSGRLLPALHAAAMIGSEMILTILNSGA 166  
Qy 159 NVNIQDAVFFPLHIAAYGHEQVTRLLKFGADVNVSGEVDRLHLASAKGFNLIAKL 218  
Db 167 NIHVTDFVHTALHCATYFQENAVRTLISASANLNLGGVANDRPILHAAAKGLTSITKL 226  
Qy 219 LMEGSKADVNAQDNEDHVPFLHFCSPFGHDIKVLQSDLEVQPHV--NTYGTPLHL 276  
Db 227 LLE--AKADPLADDEGQALHYAAKSGSLVINMLIKQVTRNDRIICARNYIGDTAHL 284  
Qy 277 ACYNGFPEVAKEIIQISGTESITKENIFSETAFHSACTYKSIDLVKFLDQNVINHQ 336  
Db 285 SCVSGRLDIVKSLDSSPTNVNVMNVFSETPLHACTGKGIELVSLMKYPGVDPNQY 344  
Qy 337 GRDHTGLHSACYGHGIRLVQFLDNGAMNLVA-----CDPSRSSGEK----- 380  
Db 345 QGDGHTALHSACYGHGIRLVQYLLENGADQQLASAFEGGALRQAGPGTNRPSKVASAI 404



C>Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 21-May-1999  
A/Accession: T04683  
R:Bevan, M.; Rose, M.; Hempel, S.; Entian, K.D.; Jesse, T.; Heijnen, L.; Vos, P.; Mewes,  
submitted to the Protein Sequence Database, July 1998  
A/Reference number: Z13381  
A/Accession: T04683  
A/Molecule type: DNA  
A/Residues: 1-553 <BEV>  
A/Cross-references: EMBL:AL031135  
A/Experimental source: Cultivar Columbia; BAC clone F8D20  
C/Genetics:  
A/Map position: 4  
A/Introns: 69/2; 107/3; 176/2; 194/3; 231/3; 289/2; 325/2; 350/1; 408/3; 440/3; 467/3; 4  
A/Note: F8D20.290

Query Match 10.8%; Score 473.5; DB 2; Length 553;  
Best Local Similarity 29.1%; Pred. No. 4.2e-18;  
Matches 134; Conservative 79; Mismatches 160; Indels 87; Gaps 15;

QY 353 IRUVQF---LLNGADMVLVADPPSSSEKDEQT-----CLMWAY 390  
Db 103 VRCVQVSTLNGSGDV-----DPSDPAYNEDAQSSYNSRLAPPTFGSSPNFALTOY 157  
QY 391 EKGH---DAIVTLKHKYKRPQDELPCNEYSP-----GGD 422  
Db 158 -KHAQDDGSAVNAQLPNSRPMEHITFTIDRPKLLSQTLSMLGELGNIQEAHAFSTAD 216  
QY 423 GSYVSV-----PSPGKTKSMT-----KEKADILLRAGLS-----HFHLQLSIE 464  
Db 217 GFSLDVFFVVDGWSQEDQPCKSQKSIITFFEHKSTNELLFACVEIPTDGTDEWEDMKQLK 276  
QY 465 FHEIIGSGFGKYKGRCKNIKVAIKRYRANTYCSKSDVDM---FCREVSLCOLNHPCV 521  
Db 277 IEKKVACSGELFRGYCSQVAKILKP-----ERNVAMELRFSQEVYINRKYRHKNV 332  
QY 522 IQFVGACLNDFPSQFAIVTOYISGGSFSLHHEQKRLDLOSKLIIADVAKGMEYLNLT 581  
Db 333 VQFAGACTRSPN-LCIVTEPTMGTSYIVFLHKGKVFQKIQSLKVALDVSKGMNYLHQ-- 389  
QY 582 QPIHRLDNLNHLLEDGHAVVADGESRFLQSLDEDNMTKQGNLRWMADEVFTQCTR 641  
Db 390 NNIIHRDLKTNILMDHEHVVVADFGVARV--QTSSGVMTAETGYRWMADEV- EHKP 446  
QY 642 YTIKADYFVALCLWEILTGEIPFAHLKPAAMAAADMAHYHIPPYGYSPKPKISSLLIRG 701  
Db 447 YDHRADVFAIVLWELLTGEIPSYLTPLOAAVGVQVQGLRKPKEPKHTELLEK 506  
QY 702 WNACPEGRPFSEVVMKLECLCNIELMSPASSNSGSL 741  
Db 507 WQDPAALRPNAETIEMNLQIREVIDLSLHKDKHGYS 546

RESULT 5  
T18276  
protein-tyrosine kinase (EC 2.7.1.112) 1 - slime mold (Dictyostelium discoideum)  
A/Cross-references: EMBL:U32174; NID:g974333; PID:g974334; PIDN:AA841125.1  
C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 17-Nov-2000  
A/Accession: T18276; A35670  
R:Nuckolls, G.H.; Osherov, N.; Loomis, W.F.; Spudich, J.A.  
Development 122, 3295-3305, 1996  
A/Title: The Dictyostelium dual-specificity kinase splA is essential for spore different  
A/Reference number: Z18852; MUID:97053827; PMID:8898241  
A/Accession: T18276  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-1584 <NUC>  
A/Cross-references: EMBL:U32174; NID:g974333; PID:g974334; PIDN:AA841125.1  
R:Tan, J.L.; Spudich, J.A.  
Mol. Cell. Biol. 10, 3578-3583, 1990  
A/Title: Developmentally regulated protein-tyrosine kinase genes in Dictyostelium discoid  
A/Reference number: A35670; MUID:90287147; PMID:1972546  
A/Accession: A35670  
A/Status: preliminary

A/Molecule type: mRNA  
A/Residues: 'R',1249-1434,'L',1436-1584 <AN>  
A/Cross-references: GB:M33785; NID:g167775; PIDN:AAA33202.1; PID:g167776  
A/Note: the authors translated the codon TAT for residue 271519 as Thr  
C/Genetics:  
A/Gene: splA  
A/Introns: 47/3; 72/2  
C/Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolc  
C/Keywords: ATP; phosphotransferase; tyrosine-specific protein kinase  
F:1287-1566/Domain: protein kinase homology <KIN>  
F:1295-1303/Region: protein kinase ATP-binding motif

Query Match 10.6%; Score 465; DB 2; Length 1584;  
Best Local Similarity 37.7%; Pred. No. 4.2e-17;  
Matches 109; Conservative 52; Mismatches 100; Indels 28; Gaps 6;

QY 454 SHFHQLQSIEFHEIIGSGFGKYKGRCKNIKVAIKRYRANTYCSKSDVDMFCREVSL 513  
Db 1280 SEVEIDFNELEFGQTIGKGFGEVGRYWRDTVAIKIYYRDQFKTKSSLVNFONEVIL 1339  
QY 514 COLNHPCVIQFVGAC-LNDPSQFAIVTOYISGGSF-----FSLHHEQKRLDLOSKLI 565  
Db 1340 SKLRHENVVQFLOACTAGGEDHHCIVTEWGGGSLRQFLTDHFNLEQNPHI-----RLK 1394  
QY 566 IAVDVAKGMEYLNLTQPIIHRDLNHNILL-----YEDGHAVVADGESRFP 612  
Db 1395 LALDIAGKMNVLHGWTPTPIIHRDLSSRNILLDHNIDPKNPVVSQRQDKICKISDFGLSR- 1453  
QY 613 LQSLDEDNMTKQGNLRWMADEVFTQCTRITKADYFVALCLWEILTGEIPFAHLKPA 672  
Db 1454 LKKEQASQMTQSGVCPYNAPFVFGKDSN-SEKSDYVYGVVLFELLTSDPEQDMKPMK 1512  
QY 673 AAADMAHYHIPPYGYSPKPKISSLLIRGNACPEGRPFSEVVMKLE 721  
Db 1513 MAHLAAVESVRPIPLTTSSKNKEILTCQWDSNPDSRPTFKQIIVHLKE 1561

RESULT 6  
B35049  
ankyrin 1, erythrocyte splice form 3 - human  
N/Alternate names: ankyrin 2.1, erythrocyte; ankyrin-R  
N/Contains: ankyrin 2.2, erythrocyte  
C/Species: Homo sapiens (man)  
C/Date: 17-Nov-1995 #sequence\_revision 17-Nov-1995 #text\_change 10-Jul-1998  
A/Accession: B35049  
R:Lambert, S.; Yu, H.; Prchal, J.T.; Lawler, J.; Ruff, P.; Speicher, D.; Cheung, M.C.; K  
Proc. Natl. Acad. Sci. U.S.A. 87, 1730-1734, 1990  
A/Title: cDNA sequence for human erythrocyte ankyrin.  
A/Reference number: A35049; MUID:90175370; PMID:1689849  
A/Accession: B35049  
A/Status: preliminary  
A/Molecule type: mRNA  
A/Residues: 1-1856 <LAM>  
C/Genetics:  
A/Gene: GDB:ANK1; ANK  
A/Cross-references: GDB:118737; OMIM:182900  
A/Map position: 9p11.2-9p11.2  
C/Superfamily: ankyrin; ankyrin repeat homology  
C/Keywords: alternative splicing  
F:2-1856/Product: ankyrin 1, erythrocyte form 3 #status predicted <MAT>  
F:2-1513,1676-1856/Product: ankyrin 2.2, erythrocyte #status predicted <NA2>  
F:44-76/Domain: ankyrin repeat homology <AN01>  
F:77-105/Domain: ankyrin repeat homology <AN02>  
F:110-142/Domain: ankyrin repeat homology <AN03>  
F:143-171/Domain: ankyrin repeat homology <AN04>  
F:172-204/Domain: ankyrin repeat homology <AN05>  
F:205-237/Domain: ankyrin repeat homology <AN06>  
F:238-270/Domain: ankyrin repeat homology <AN07>  
F:271-303/Domain: ankyrin repeat homology <AN08>  
F:304-336/Domain: ankyrin repeat homology <AN09>  
F:337-369/Domain: ankyrin repeat homology <AN10>  
F:370-402/Domain: ankyrin repeat homology <AN11>  
F:403-435/Domain: ankyrin repeat homology <AN12>







A;Gene: ATSP.F2401.13  
A;Map position: 1  
A;Introns: 149/3; 301/3  
C;Superfamily: Kinase-related transforming protein; protein kinase homology

Query Match 10.1%; Score 444.5; DB 2; Length 390;  
Best Local Similarity 31.6%; Pred. No. 9.8e-17;  
Matches 120; Conservative 74; Mismatches 129; Indels 57; Gaps 13;

QY 411 LPCNEYSQGGGGSVSPPLGKTKSMTEKAD-----ILLRA-----GLPSH 455  
Db 28 LFCNPESS-----SKTNRYAEAEETMEKFRFDSMESWSMILESENVTWEASKGREE 78

QY 456 FHLQLSEIEFBHIIIGSGFGKYGKRCNKIIVAKRYRANTYCKSDVDM---PCREVSI 512  
Db 79 WTADLSQLFIGNKFPASGAHSRIYGIYQKQAVAVKVPRIPTKKEKTRAKLQKQKSEVAL 138

QY 513 LCQLNHPCVIQVAGCLNDPSPQFAIVQYISGGSLFSLHBEQKRI-LDLOSLLIIVDVA 571  
Db 139 LGRLEPHNIVQFIAAC-KKPPVYCIITEYMSQGNLRMLNKKPEYSLSIETVLEALDIS 197

QY 572 KGWELVNLNTOPIIHRDLNSHNLIVEDGHVAVADFGESRFLQSLDENMTKQPGNL--- 628  
Db 198 RGMEXYDH--SQGVHRLDKSNLLNDEMVRVKVADFGISCL-----ETQCREAKGNMGTY 250

QY 629 RWAPEVFTQCTRYTIKADVFYSALCLMEILTGEIFFAHLKPAAAAADMAVHHIRPPIGY 688  
Db 251 RWAPEMIKE-KEYTRKVDVYSGIVLMEILTALLPFPQGMTPVQAAFAVAEKNRPPLPA 309

QY 689 STPKPISSLLIRGWACPGREPERFEVWKLE---ECLC-NIELMSPASSNSSSL----- 740  
Db 310 SCQPALAHLIKRCWSENPSKPDFGSLNIVAVLEKYDECYKGLPLTSHASLTTKTKAILDH 369

QY 741 -----SPSSSSDCLVN 751  
Db 370 LKGCVTISISPPSSSSVPVN 389

RESULT 12  
S37771  
ankyrin, erythrocyte - mouse  
C;Species: Mus musculus (house mouse)  
C;Date: 27-May-1994 #sequence\_revision 03-Aug-1995 #text\_change 13-Aug-1999  
C;Accession: S37771  
R;Birkenmeier, C.S.; White, R.A.; Peters, L.L.; Hall, E.J.; Lux, S.E.; Barker, J.E.  
J. Biol. Chem. 268, 9533-9540, 1993  
A;Title: Complex patterns of sequence variation and multiple 5' and 3' ends are found an  
A;Reference number: S37771; MUID:93252825; PMID:8486643  
A;Accession: S37771  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-1848 <BIR>  
A;Cross-references: EMBL:X69063; NID:g311816; PIDN:CAA48801.1; PID:g311817  
C;Superfamily: ankyrin; ankyrin repeat homology  
C;Keywords: alternative splicing  
F;48-80/Domain: ankyrin repeat homology <AN01>  
F;81-113/Domain: ankyrin repeat homology <AN02>  
F;114-146/Domain: ankyrin repeat homology <AN03>  
F;147-175/Domain: ankyrin repeat homology <AN04>  
F;176-208/Domain: ankyrin repeat homology <AN05>  
F;209-241/Domain: ankyrin repeat homology <AN06>  
F;242-274/Domain: ankyrin repeat homology <AN07>  
F;275-307/Domain: ankyrin repeat homology <AN08>  
F;308-340/Domain: ankyrin repeat homology <AN09>  
F;341-373/Domain: ankyrin repeat homology <AN10>  
F;374-406/Domain: ankyrin repeat homology <AN11>  
F;407-439/Domain: ankyrin repeat homology <AN12>  
F;440-472/Domain: ankyrin repeat homology <AN13>  
F;473-505/Domain: ankyrin repeat homology <AN14>  
F;506-538/Domain: ankyrin repeat homology <AN15>  
F;539-571/Domain: ankyrin repeat homology <AN16>  
F;572-604/Domain: ankyrin repeat homology <AN17>  
F;605-637/Domain: ankyrin repeat homology <AN18>

F;638-670/Domain: ankyrin repeat homology <AN19>  
F;671-703/Domain: ankyrin repeat homology <AN20>  
F;704-736/Domain: ankyrin repeat homology <AN21>  
F;737-769/Domain: ankyrin repeat homology <AN22>  
F;770-802/Domain: ankyrin repeat homology <AN23>

Query Match 10.1%; Score 444; DB 2; Length 1848;  
Best Local Similarity 28.5%; Pred. No. 6.7e-16;  
Matches 142; Conservative 63; Mismatches 159; Indels 134; Gaps 15;

QY 60 LNYRTE-----NGLSLHLCCICGGKKSHTRLMLKGLRPSRLTRNGFTALHLAVYKDN 113  
Db 329 LOYNAEIDITLDLHDHTPLHVAACHGHRV-AKVLLDKGAKNSALNGFTPLHIACCKNH 387

QY 114 ASLITSLHSGADIQQVGYGGTALHIATTAGHLEAADVLLQHGANYNIODAVPFTPLHI 173  
Db 388 IRVMBELLKTGASIDAVTESGLTPLHVASFWGHLPIVKNLLQRCASPNVSNVKVETELHM 447

QY 174 AAYYHEQVTRILLKFGADVNVSGVDRPLHLASAKGFLNIAKLMBEESKADVNAQDN 233  
Db 448 AARAGHTEVAKYLLQNKAKAKAKADQDTPLHCAARIGHTGMVXLLLENG--ASPNIATT 505

QY 234 EDHVPLHFCSRFGH-----HDIVKY-----LLQSD-- 258  
Db 506 AGHTPLHTAAREGHVDLTALLEKEASOACTKKGFTPLHVAAYGKVRLAELLEDHDAH 565

QY 259 -----LEV-----QPHVNVNYGDTPLHLACYNGKPEVAKEI 289  
Db 566 PNAAGKNGLTPLHVAHVHNNLDIVKLLPRGSGPHSPAWNGYTPPLHIAAKQNIQEVARS 625

QY 290 IQISGTESLTENISETAFHSACTYKSIDLVKFLDQNV----- 330  
Db 626 LOYGG--SANAESVOGVTPHLHAAQEGHT-EMVALLLSKQANGNLGNKSGLTPLHLVSOE 682

QY 331 -----ININHQGRDGHGTGLHSACYHGHIRLVQFLDNGADMNLVACDPSRSS 377  
Db 683 GHVPVADVLIKHGVTVDATTRMGYTPLVHVAHYGNIKLVKELLQHQADV-----A 733

QY 378 GEKDRQOTCLMWAYEKGHDAIVT-LLKHYKRPQDELPCEHYEQPGDGG-----SYVSV 429  
Db 734 KTKLGYSPHLQAAQOQGTDTIVTLKNGASP-----NEVSSNGTTPLAIAKRLGYISVT 787

QY 430 SPLGKIKSMTEKADILL 447  
Db 788 DVL---KVTDTSTVVLV 802

RESULT 13  
T13940  
ankyrin - fruit fly (Drosophila melanogaster)  
C;Species: Drosophila melanogaster  
C;Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 17-Nov-2000  
C;Accession: T13940  
R;Dubreuil, R.R.; Yu, J.  
Proc. Natl. Acad. Sci. U.S.A. 91, 10285-10289, 1994  
A;Title: Ankyrin and beta-spectrin accumulate independently of alpha-spectrin in Drosoph  
A;Reference number: Z17820; MUID:95024098; PMID:7937942  
A;Accession: T13940  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-1549 <DUB>  
A;Cross-references: EMBL:L35601; NID:g557083; PID:g557084; PIDN:AAC37208.1  
C;Genetics:  
A;Cross-references: FlyBase:FBgn0011747

Query Match 10.1%; Score 443; DB 2; Length 1549;  
Best Local Similarity 25.2%; Pred. No. 6.1e-16;  
Matches 179; Conservative 98; Mismatches 217; Indels 216; Gaps 29;

QY 57 KVNLYRTEGSLSLHLCCICGGKKSHTRLML-----KGLRPSRL----- 97  
Db 389 KANPNARALNGTPLHIAC----KKNRMVLELLIKHGKANGIGATTESGLTPLHVASFWGC 444

```

QY 98 -----TRNGFTALHLAVYKDNALITSLHSGADIQQVGVGGLTALHI 140
Db 445 INIVYLLQHEASADLTFTIGETPLHLAARANQADRIILRLS-AKVDAIVREGQTPHLV 503
QY 141 ATIAHLEAADVLLQGANVNIQDAVFFTLHIAAYVGHQVTRLLKFGADVNSGEVG 200
Db 504 ASRLGNINIIMLLQGAENIQAQNDKYSAHIAAKEGQENIVQVLLLENGAENNAVTXG 563
QY 201 DRPVLASAKFNLIAKLMEEGSKADVNAQDNEDHVPFLHFCRFGHDIIVKYLQSDLE 260
Db 564 FTPLHLACKYQKNVQQLQNG--ASIDFGKNDVTPLHVAHYNNPISIVELLKNG-- 619
QY 261 VQPHVNIYGTPLHLACYNGKPEVAKBIIQIGSTESLTAKENIFSETAF---HSACTYK 317
Db 620 SSPNLCAKNGQCAHIAKKNYLEIAQLQ-HGAD---VNIISKSGFSLHLA-AQGG 673
QY 318 SIDLVKFLDQVNI-----LPSHFLQLSEIEFHEIIGSGFGKYVKGCRNKIVA-IKRYR 493
Db 674 NDMVQLLEYGVISAAXKXGLTPLHVAQEGHVLVSQILLHEGANISERTNGVTPHLM 733
QY 347 ACYHGHIRLVQFLDNGADMLVACDPSRSGSEKDEOTCLMWAYEKHDAIVT-LLKHVK 405
Db 734 AAHVGHLDLVKFTENDADIEM-----SSNIG---YTPLHQAQOGHIMIINLLRHKA 784
QY 406 RPQDELPCNEYSQPGDQ-----SYVSVPSPL-----GKIKSMYKE 441
Db 785 NP-----NALTQDGNALHIAASNLGYVTVMESLKIIVTSVINSIGAIEBKLVMTPE 838
QY 442 KADILLARAG-----LPSHFLQLSEIEFHEIIGSGFGKYVKGCRNKIVA-IKRYR 493
Db 839 LMQTLLSDSDSCDILLDHNH-----YKMATDLDKANYGQDQ 878
QY 494 ANTYCSKSDVDMFCREYSILCOLNHPCVIQVQACIL-----NDPSQFAIVTYISGSL 547
Db 879 KNFDTTNDHDL--TDVSV--LNKKEILPNESCIELTEIGHKPDNVVIAASQVHLGFL 933
QY 548 FSLHLEQKRLDLSKLIIVDVAKG--MEYLNLTQPII-----HR 587
Db 934 VSFL-----VDARGSGMRGYRNGRIIVPPKACAEPTITCRYVYKQQR 977
QY 588 DLNHSNILLYEDGHAVVADEGSRFLQSLDEDNMTKQ-----GNLR 629
Db 978 VVNPPPLM---EGEALV-----SRILEMSPVQGMFLSPITLEVPHVGTLR 1019

RESULT 14
S29851
protein kinase 6 (EC 2.7.1.1) - soybean
C:Species: Glycine max (soybean)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 24-Sep-1999
C:Accession: S29851; S27760
R:Feng, X.H.; Zhao, Y.; Bottino, P.J.; Kung, S.
Biochim. Biophys. Acta 1172, 200-204, 1993
A:Title: Cloning and characterization of a novel member of protein kinase family from so
C:Keywords: ATP; phosphotransferase
A:Reference number: S29851; MUID:93176812; PMID:8439562
A:Accession: S29851
A:Molecule type: mRNA
A:Residues: 1-462 <FEN>
A:Cross-references: EMBL:M67449; NID:g170046; PID:AAA34002.1; PID:g170047
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo
F:154-419/Domain: protein kinase homology <KIN>

Query Match 10.0%; Score 439.5; DB 2; Length 462;
Best Local Similarity 33.6%; Pred. No. 2.2e-16;
Matches 112; Conservative 66; Mismatches 128; Indels 27; Gaps 9;

QY 430 SPLGKIKSMY-KKADILLRAG-----LPSHFLQLSEIEFHEIIGSGFGKYVKG 480
Db 114 SPNRQVSKNRKDSAWTKLLDNGGKITAVETAENWVMSQLFGLKFAHGAHSRLYHG 173
QY 481 RCNKKIVAIK-----RYRANTYCSKSDVDMFCREYSILCOLNHPCVIQVVGACLDPSQF 535

```

```

Db 174 VYKDEAVAVKIIVVPEDDGNALASRLKQFIREVTLLSRLHQNVIKFSAC-RKPPVY 232
QY 536 AIYTVIYSGSLFSLIHE-QKRILDLSKLIITAVDVAKGMEYLNLTQPIIHRDLNSHI 594
Db 233 CIITEYLAESLRLYHLEHQTISLQKLIATFALDIARGMEYIH--SGGVHRRDLKPENI 290
QY 595 LLYEDGHAVVADGEGSRFLQSLDEDNMTKQPNLWMAPEVFTQCTRYTIKADIVFSVLC 654
Db 291 LINEDNHLKXADFGIA--CEASCDLLADDPTYRWMAPEMIKR-KSYGKXVDVYSGLI 347
QY 655 LWELLGETIPFAHLKFAAADAAMVHHIRPPYIGYIPKPISSLLIRGNWACPGRPERSE 714
Db 348 LWEMLTGITPYEDMNPQAAFAVAVKNSRPPIFSPCPPAMRALIEQCSLQDPKRPEFWQ 407
QY 715 VVMKLECLNLTMLSPASSNSSLSPSSSSD 747
Db 408 VKVLE-----QFESSLASDGTLSLVNPNCD 434

RESULT 15
T10671
protein kinase homolog F6E21.90 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 19-May-2000
C:Accession: T10671
R:Bevan, M.; Lennard, N.; Quail, M.; Harris, B.; Rajandream, M.A.; Barrell, B.G.; Banc
submitted to the Protein Sequence Database, June 1999
A:Reference number: Z16533
A:Accession: T10671
A:Molecule type: DNA
A:Residues: 1-412 <SEV>
A:Cross-references: EMBL:AL049914; GSPDB:GN00062; ATSP:F6E21.90
A:Experimental source: cultivar Columbia; BAC clone F6E21
C:Genetics:
A:Gene: ATSP:F6E21.90
A:Map position: 4
A:Introns: 300/2
C:Superfamily: kinase-related transforming protein; protein kinase homology
F:135-392/Domain: protein kinase homology <KIN>

Query Match 10.0%; Score 438; DB 2; Length 412;
Best Local Similarity 30.4%; Pred. No. 2.3e-16;
Matches 127; Conservative 71; Mismatches 160; Indels 60; Gaps 16;

QY 334 NHQGRDCHTGLHSACHYGHIRLVQFLDNGADMLVACDPSRSGEKD-----BQT 384
Db 14 NHNNNNYAFYQDFY-----QKLGEETNMSYDSMQTSNAGSGVSMVSDNSVSGSSD 66
QY 385 CLMWAYEKGHDAIVTLKHYKRPQDELPCNEYSQPGDGSYVSPSLGKIKSMYKEKA- 443
Db 67 ALI-----GHPG-----LKMRRP-----YSL--SDGQSVFRP--GKTHALNDAL 104
QY 444 -----DILLRAGLPSH--FHIQLSEIEFHEIIGSGFGKYVKGCRNKIVAIK---RYR 493
Db 105 AQALMDSKYPTREGLVNVEWTTDLRLKLMGPAFAQAGFKLYRGTYNGEDVAIKLLERSD 164
QY 494 ANTYCSKSDVDMFCREYSILCOLNHPCVIQVVGACLDNPSQFAIVTYISGSLFSL-H 552
Db 165 SNPEKAQALEQQQVEVSMFLAKEPNIVRFIGACIK-PMWCVITEYAKSGSVQRFETK 223
QY 553 EQKRILDLSKLIITAVDVAKGMEYLNLTQPIIHRDLNSHILYEDGHAVVADFGSRF 612
Db 224 RQNRVPLKLVQALDVARGMAYVHE--RNFIHRDLKSDNLLISADRSIKIADFGVARI 281
QY 613 LQSLDEDNMTKQPNLWMAPEVFTQCTRYTIKADIVFSVLCLEILTGETIPFAHLKPA 672
Db 282 --EVQTEGMPETGTGRWMAPEMI--OHRPYTKQVDVVSFGIVLWELLITGLLPQNMTAVQ 338
QY 673 AAADMAHYHRRPPIGYISPKPISSLLIRGNWACPGRPERFSEVVMKLE----ECLNI 726
Db 339 AAFVAVNRGVRTPVADCLPVLGEIMTRCWDADPEVRPCFAEIVNLEAAETEIMTVN 396

```

Search completed: September 2, 2004, 14:25:02  
Job time : 57 secs

---

This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 5, 2004, 23:38:08 ; Search time 5035.83 Seconds  
(without alignments)  
17938.068 Million cell updates/sec

Title: US-10-626-173-1

Perfect score: 3025

Sequence: 1 gtcgaccacagctccggcc.....aaaaaaaaaaggcgccgc 3025

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

```
EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	734.6	24.3	876	14	CF551822
2	721	23.8	756	12	AGENCOURT
3	702	23.2	757	12	BG219957
4	682.8	22.6	813	12	BG209217

#### ALIGNMENTS

```
RESULT 1
CF551822
LOCUS      AGENCOURT 15595443 NIH MGC 183 Homo sapiens cDNA clone
DEFINITION IMAGE:30530169 5', mRNA sequence.
ACCESSION  CF551822
VERSION     CF551822.1 GI:34888656
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 876)
NTH-MGC http://mgc.nci.nih.gov/.
AUTHORS    National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE      Unpublished (1999)
JOURNAL
COMMENT     Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgabs-r@mail.nih.gov
            Tissue Procurement: Dr. Michael Brownstein
            CDNA Library Preparation: Invitrogen Corp
            CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
```

5	677.4	22.4	793	12	BG216613
6	665	22.0	665	13	EX501623
7	618.6	20.4	902	10	BE897696
8	617.8	20.4	656	13	EX502807
9	598.8	19.6	617	12	BG213152
10	591.8	19.6	617	12	BG216471
11	590.8	19.5	785	12	BG215938
12	589.4	19.5	613	12	BG212638
13	588.2	19.4	629	12	BG186716
14	587.4	19.4	638	9	AL599654
15	587.2	19.4	785	12	BG181490
16	587	19.4	616	12	BG182504
17	586.8	19.4	632	12	BG210010
18	586.2	19.4	616	12	BG187210
19	584	19.3	607	12	BG214260
20	583.4	19.3	604	12	BG221417
21	582.6	19.3	619	12	BG215301
22	582.6	19.3	638	12	BG185641
23	582.2	19.2	616	12	BG218348
24	582	19.2	614	12	BG185642
25	577.2	19.1	618	12	BG215300
26	576.8	19.1	604	12	BG186114
27	576.8	19.1	617	12	BG183556
28	575.6	19.0	614	12	BG198554
29	574	19.0	617	12	BG188258
30	572.8	18.9	617	12	BG204728
31	572.8	18.9	634	12	BG215865
32	571.2	18.9	615	12	BG196050
33	570.6	18.9	616	12	BG190428
34	570.4	18.9	620	12	BG213672
35	568.4	18.8	889	12	BG190429
36	565.8	18.7	612	12	BG196530
37	562.8	18.6	605	12	BG216472
38	558.6	18.5	602	12	BG192433
39	558.4	18.5	625	12	BG219958
40	555.4	18.4	603	12	BG187211
41	552.8	18.3	614	12	BG211040
42	547.8	18.1	609	12	BG205918
43	547	18.1	626	12	BG188259
44	540.8	17.9	892	12	BG202691
45	537.8	17.8	611	12	BG210009

CF551822 876 bp mRNA linear EST 22-SEP-2003

AGENCOURT 15595443 NIH MGC 183 Homo sapiens cDNA clone

IMAGE:30530169 5', mRNA sequence.

CF551822

CF551822.1 GI:34888656

EST.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 876)

NTH-MGC http://mgc.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Daniela S. Gerhard, Ph.D.

Office of Cancer Genomics

National Cancer Institute / NIH

Bldg. 31 Rm10A07 Bethesda, MD 20892

Email: cgabs-r@mail.nih.gov

Tissue Procurement: Dr. Michael Brownstein

CDNA Library Preparation: Invitrogen Corp

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov  
Plate: NDAM618 row: 0 column: 10  
High quality sequence start: 15  
High quality sequence stop: 670.  
Location/Qualifiers  
1..876

## FEATURES

source

/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:30530169"  
/lab\_host="DH10B-TonA (T1 and T5 phage resistant)"  
/clone\_lib="NIH\_MGC\_183"  
/note="Organ: Pooled muscle (cardiac and skeletal);  
Vector: pCMV-SPORT6.1; Site 1: EcoRV (destroyed); Site 2:  
NotI; Library is oligo-dT primed and directionally cloned  
(EcoRV site is destroyed upon cloning). Average insert  
size 1.7. Library was constructed by Invitrogen."

## ORIGIN

Query Match 24.3%; Score 734.6; DB 14; Length 876;  
Best Local Similarity 97.6%; Pred. NO. 2.7e-126;  
Matches 830; Conservative 0; Mismatches 10; Indels 10; Gaps 8;  
399 ACTTCTGCTTCAGTGGAGCTGATATACAGCAGGTGGATACGGTGGCCCTCCTGCTGC 458  
Db |||||  
28 ACTTCTGCTTC-CAGTGGAGCTGATATACAGCAGGTGGATACGGTGGCCCTCCTGCTGC 86  
Qy |||||  
459 CTCATATTGCTACAATAGCTGGCCACCTAGAGGCTGCTGATGCTGTTTGGCAACATGGA 518  
Db |||||  
87 CTCATATTGCTACAATAGCTGGCCACCTAGAGGCTGCTGATGCTGTTTGGCAACATGGA 146  
Qy |||||  
519 GCTATGTCATATCAAGTGCAGTCTTTTCTACTCCATGCAATATGACGGTACTAT 578  
Db |||||  
147 GCTATGTCATATCAAGTGCAGTCTTTTCTACTCCATGCAATATGACGGTACTAT 206  
Qy |||||  
579 GGACATGAACAGTAACTCGGCTCTTTTGAATTTGGTCTGATGTAATGTAAGTGGT 638  
Db |||||  
207 GGACATGAACAGTAACTCGGCTCTTTTGAATTTGGTCTGATGTAATGTAAGTGGT 266  
Qy |||||  
639 GAAGTTGAGATAGACCCCTCCACCTAGCATCTGCAAAAGGATCTTGAATATGCAAAA 698  
Db |||||  
267 GAAGTTGAGATAGACCCCTCCACCTAGCATCTGCAAAAGGATCTTGAATATGCAAAA 326  
Qy |||||  
699 CTCTTGATGGAAGGAGCAGACAGATGTGATGCTCAAGATATGACGACCATGTC 758  
Db |||||  
327 CTCTTGATGGAAGGAGCAGACAGATGTGATGCTCAAGATATGACGACCATGTC 386  
Qy |||||  
759 CCATCCATTTCTGTTCTCGATTTGGACACCATGATATAGTTAAGTATCTGCTGCAAGT 818  
Db |||||  
387 CCATCCATTTCTGTTCTCGATTTGGACACCATGATATAGTTAAGTATCTGCTGCAAGT 446  
Qy |||||  
819 GATTTGGAAGTCAACCTCATGTTGTAATATCTATGGAGATACCCCTTACACCTGGCA 878  
Db |||||  
447 GATTTGGAAGTCAACCTCATGTTGTAATATCTATGGAGATACCCCTTACACCTGGCA 506  
Qy |||||  
879 TGCTACATGTCACAAATTTGAAGTTGCCAAGAAATCATCCAAATATCAGAAACAGAAAGT 938  
Db |||||  
507 TGCTACATGTCACAAATTTGAAGTTGCCAAGAAATCATCCAAATATCAGAAACAGAAAGT 566  
Qy |||||  
939 CTGACTAAGGAAACATCTTCAGTGAACAGCTTTTTCATAGTCTGTTGTAACCTATGGCAAG 998  
Db |||||  
567 CTGACTAAGGAAACATCTTCAGTGAACAGCTTTTTCATAGTCTGTTGTAACCTATGGCAAG 626  
Qy |||||  
999 AGCATTCACCTAGTCAAAATTTCTCTTGATCAGAATGTCTATAACATCAACACCAAGG- 1057  
Db |||||  
627 AGCATTCACCTAGTCAAAATTTCTCTTGATCAGAATGTCTATAACATCAACACCAAGG 686  
Qy |||||  
1058 AAGGGATGGGACACT-GGATTAACATCT-GCTTGCTACCAAGGTCAACATTCG-CTCGT 1114  
Db |||||  
687 AAGGGATGGGACACTINGGATTAACATCTINGCTTGTACCAAGGTCAACATTCGCTGCT 746  
Qy |||||  
1115 TCAGTTCTTACTGTGATAATGGAGCTGATATGATCTAGT-GGCTTGTGATCCACAGG- 1172  
Db |||||

Db 747 TCAGTTCTTACTGTGATATGGAGCTGATATGAATCTAGTGGGCTTCTGATCCACGACGN 806  
Qy 1173 TCTAGTGGTGAAGAGATGACGACACATGTTT---GATGGGCTTATGAAAAAGGCAT 1229  
Db |||||  
807 TCTAGTGGTGAAGAGATGACGACACATGTTTGGTGGGCTTATGAAAAAGGCAT 866  
Qy 1230 GATGCCATTG 1239  
Db |||||  
867 GATGCCATTG 876  
Bg227974 756 bp mRNA linear EST 21-APR-2001  
R5741894 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.  
Bg227974  
Bg227974.1 GI:13748097  
EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
MAMMALIA; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE  
AUTHORS Harrington, J.J., Sherf, B., Rundlett, S., Jackson, P.D., Perry, R.,  
Cain, S., Leventhal, C., Thornton, M., Ramachandran, R.,  
Whittington, J., Lerner, L., Costanzo, D., McElligott, K., Booser, S.,  
Mays, R., Smith, E., Veloso, N., Klika, A., Hess, J., Cothren, K., Lo, K.,  
Offenbacher, J., Danzig, J., and Ducar, M.  
TITLE Creation of genome-wide protein expression libraries using random  
activation of gene expression  
JOURNAL Nat. Biotechnol. 19 (5), 440-445 (2001)  
MEDLINE 21227151  
PUBMED 11329013  
COMMENT Contact: Scott J. Cain  
Athersys, Inc.  
3201 Carnegie Ave, Cleveland, OH 44115, USA  
Tel: 216 431 9900  
Fax: 216 361 9596  
Email: scain@athersys.com  
High quality sequence stop: 520.  
Location/Qualifiers  
1..756  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone\_lib="Athersys RAGE Library"  
/notes="See 'Creation of Genome-wide Protein Expression  
Libraries using Random Activation of Gene Expression',  
Nature Biotechnology, in press. Note that even though the  
cell type indicated is HT1080, since a random activation  
method was used, these sequence tags are not necessarily  
expressed in HT1080 under normal circumstances."

Query Match 23.8%; Score 721; DB 12; Length 756;  
Best Local Similarity 97.3%; Pred. No. 9.6e-124;  
Matches 733; Conservative 0; Mismatches 20; Indels 0; Gaps 0;  
593 AACTCGCCTTCTTTGAAATTTGGTCTGATGTAATTAAGTGGTGAAGTTGGAGATAG 652  
Db 1 AACTCGCCTTCTTTGAAATTTGGTCTGATGTAATTAAGTGGTGAAGTTGGAGATAG 60  
Qy 653 ACCCTCCACCTAGCATCTGCAAAAGGATTTCTGATATTCGAAACCTCTGATGGAAGA 712  
Db 61 ACCCTCCACCTAGCATCTGCAAAAGGATTTCTGATATTCGAAACCTCTGATGGAAGA 120  
Qy 713 AGCAGCAAGAGATGGAATGCTCAAGATAATGAAGACCATGTGCCACTCCATTTCTG 772  
Db 121 AGCAGCAAGAGATGGAATGCTCAAGATAATGAAGACCATGTGCCACTCCATTTCTG 180  
Qy 773 TTCTCGATTTGACACCATGATATGATTAAGTATCTGCTCAAGATGATTTGGAAGTCA 832  
Db |||||





```
QY 593 AACTCGCTCTCTTTTGAATTTGGTCTCATGTAAATGTAAAGTGGTGAAGTGGAGATAG 652
Db |||||
QY 653 ACCCTCCACCTAGACTCTGCAAAAGGATCTTGTAATATTCGAAATCTCTTGATGGAAGA 712
Db |||||
QY 61 ACCCTCCACCTAGACTCTGCAAAAGGATCTTGTAATATTCGAAATCTCTTGATGGAAGA 120
QY 713 AGCAGCAAGAGAGATGTAAGTCTCAAGATATGAAGACCATGTCCTCACTCCATCTTCG 772
Db |||||
QY 121 AGCAGCAAGAGAGATGTAAGTCTCAAGATATGAAGACCATGTCCTCACTCCATCTTCG 180
QY 773 TTCTCGATTGGACACCATGATATAGTAAATCTGCTGCAAAAGTGAATTTGGAAGTTCA 832
Db |||||
QY 181 TTCTCGATTGGATACCATGATATAGTAAATCTGCTGCAAAAGTGAATTTGGAAGTTCA 240
QY 833 ACCTCATGTTGTAATATCTATGGAGATACCCCTTACACCTGGCATGCTCAATGGCAA 892
Db |||||
QY 241 ACCTCATGTTGTAATATCTATGGAGATACCCCTTACACCTGGCATGCTCAATGGCAA 300
QY 893 ATTGGAAGTTGCCAAGGAAATCATCCAAATATCAGGAACAGAAAGTCTGACTAAGGMAAA 952
Db |||||
QY 301 ATTGGAAGTTGCCAAGGAAATCATCCAAATATCAGGAACAGAAAGTCTGACTAAGGMAAA 360
QY 953 CATCTTCAGTGAACAGCTTTTCATAGTCTGTACCTATGCGCAAGAGATTCACCTAGT 1012
Db |||||
QY 361 CATCTTCAGTGAACAGCTTTTCATAGTCTGTACCTATGCGCAAGAGATTCACCTAGT 420
QY 1013 CAAATTTCTTTGATCAGAAATGTCATAAATCATCAACCAAGCAAGGATGGSCACAC 1072
Db |||||
QY 421 CAAATTTCTTTGATCAGAAATGTCATAAATCATCAACCAAGCAAGGATGGSCACAC 480
QY 1073 TGAATTAACACTCTGCTTACACAGGTCACATTCGCTGTTTCAAGTCTTACTGGATAA 1132
Db |||||
QY 481 TGAATTAACACTCTGCTTACACAGGTCACATTCGCTGTTTCAAGTCTTACTGGATAA 540
QY 1133 TGAGCTGATATGAATCTAGTGGCTGTGATCCGAGAGTCTAGTGGTGAAGAAGATGA 1192
Db |||||
QY 541 TGAGCTGATATGAATCTAGTGGCTGTGATCCGAGAGTCTAGTGGTGAAGAAGATGA 600
QY 1193 GCAGACATGTTGATGGGCTTTATGAAAAGGGCATGATGCAATTTGCACACTCTCTGAA 1252
Db |||||
QY 601 CCAGACATGTTGATGGGCTTTATGAAAAGGGCATGATGCAATTTGCACACTCTCTGAA 658
QY 1253 GCATTAAGAGACCAAGATGAATGGCTGTGATGATGATGATGATGATGATGATGATGAT 1312
Db |||||
QY 659 GC-TTAAAGAGACC-CCAGATGAATGGCCCTG-AATGAATATTTCTCAACCT-GAAGAGA 714
QY 1313 TGGCTCTATGCTGCTGTTCCATCACCTTGGGGAAGATTTAAAGCATGACAAAAGAGAA 1372
Db |||||
QY 715 TGGCTCTATGCTGCTGTTCCATCACCTTGGGGAAGATTTAAAGCATGACAAAAGAGAA 1372
Db |||||
QY 1373 GGCAGATATTCCTCTC 1388
Db |||||
QY 774 GGCCGATTTTCTCTC 789
```

```
RESULT 6
BX501623 665 bp mRNA linear EST 04-SEP-2003
LOCUS DXFZp779A1164 r1 779 (synonym: hnccl) Homo sapiens cDNA clone
DEFINITION DXFZp779A1164 5', mRNA sequence.
ACCESSION BX501623
VERSION BX501623.1 GI:32022118
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 665)
Wambutt,R., Heubner,D., Mewes,H.W., Weil,B., Amid,C., Osanger,A.,
Fobo,G., Han,M. and Wiemann,S.
```

```
EST (Wambutt,R., Heubner,D., Mewes,H.W., Weil,B., Amid,C., et al.)
Unpublished (2003)
Contact: MIPS
MIPS
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by AGOWA (Berlin/Germany) within the cDNA sequencing
consortium of the German Genome Project.
No si sequence available.
This clone (DXFZp779A1164) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
```

```
FEATURES
Location/Qualifiers
1..665
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DXFZp779A1164"
/tissue_type="liver"
/dev_stage="fetal"
/lab_host="DH10B"
/clone_lib="779 (synonym: hnccl)"
/note="Vector: pSport1_Sfi; Site_1: SfiIA; Site_2: SfiIB"
```

```
ORIGIN
Query Match 22.0%; Score 665; DB 13; Length 665;
Best Local Similarity 100.0%; Pred. No. 2.5e-113;
Matches 665; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 573 TACTATGGACATGAACAGGTAACTCGCTCTCTTTTGAATTTGGTCTGATGTAATGTA 632
Db 1 TACTATGGACATGAACAGGTAACTCGCTCTCTTTTGAATTTGGTCTGATGTAATGTA 60
QY 633 AGTGGTGAAGTTGGAGATAGACCCCTCCACCTAGCATCTGCAAAAGGATCTTCAATATT 692
Db 61 AGTGGTGAAGTTGGAGATAGACCCCTCCACCTAGCATCTGCAAAAGGATCTTCAATATT 120
QY 693 GCAAACTCTTGATGGAAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 752
Db 121 GCAAACTCTTGATGGAAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
QY 753 CATGCCCCACTCCATCTTCGTTCTCGATTTGGACACCATGATATAGTTAAGTATCTGCTG 812
Db 181 CATGCCCCACTCCATCTTCGTTCTCGATTTGGACACCATGATATAGTTAAGTATCTGCTG 240
QY 813 CAAAGTGATTTGGAAGTTCAACCTCATGTTGTTAATATCTATGGAGATACCCCTTACAC 872
Db 241 CAAAGTGATTTGGAAGTTCAACCTCATGTTGTTAATATCTATGGAGATACCCCTTACAC 300
QY 873 CTGGCATGTACAATGGCAAAATTTGAAGTTCGCAAGGAAATCATCCAAATATCAGGAGAA 932
Db 301 CTGGCATGTACAATGGCAAAATTTGAAGTTCGCAAGGAAATCATCCAAATATCAGGAGAA 360
QY 933 GAAAGTCTGACTAAGGAAACATCTTCAGTGAACAGCTTTTCATAGTCTGTTGACCTAT 992
Db 361 GAAAGTCTGACTAAGGAAACATCTTCAGTGAACAGCTTTTCATAGTCTGTTGACCTAT 420
QY 993 GGCAGAGCAITTGACCTAGTCAAAATTTCTTGTATCAGATGTCATAAATCAACAC 1052
Db 421 GGCAGAGCAITTGACCTAGTCAAAATTTCTTGTATCAGATGTCATAAATCAACAC 480
QY 1053 CAAGGAAGGATGGGACACTGATTAACATCTGCTTCTACCGGTCACATTCGCTG 1112
Db 481 CAAGGAAGGATGGGACACTGATTAACATCTGCTTCTACCGGTCACATTCGCTG 540
QY 1113 GTTCACTTCTTACTGGATAATGGAGCTGATATGAATCTAGTGGCTTGTGATCCAGCAGG 1172
Db 541 GTTCACTTCTTACTGGATAATGGAGCTGATATGAATCTAGTGGCTTGTGATCCAGCAGG 600
QY 1173 TCTAGTGGTGAAGAGATGAGCAGACATGTTGATGCTGGCTTATGAAAGGCGCATCAT 1232
Db |||||
```

```

Db      601 TCTAGTGGTGAAGAAGATGAGCAGACATGTTTCATGTGGGCTTATGAAAAAGGGCATGAT 660
QY      1233 GCCAT 1237
        |||||
Db      661 GCCAT 665

RESULT 7
BB997696
LOCUS   601439226F1 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:3924055 5',
DEFINITION
mRNA sequence.
ACCESSION BB997696
VERSION   BB997696.1 GI:10363419
KEYWORDS EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 902)
AUTHORS   NIH-MGC http://mgi.nci.nih.gov/.
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
          Email: cgapbs-remail.nih.gov
          Tissue Procurement: ATCC/DCTD/DTF
          cDNA Library Preparation: Life Technologies, Inc.
          cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
          DNA Sequencing by: Incyte Genomics, Inc.
          Clone Distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LLNL at:
          http://image.llnl.gov
          Plate: L1AM9761 row: d column: 08
          High quality sequence stop: 689.

FEATURES             source
    source
        1..902
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:3924055"
            /tissue_type="melanotic melanoma"
            /lab_host="DH10B (phage-resistant)"
            /clone_lib="NIH MGC_72"
            /note="Organ: skin; Vector: pCMV-Sport6; Site_1: NotI;
            Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
            Average insert size 2 kb. Library constructed by Life
            Technologies."

ORIGIN
Query Match      20.4%; Score 618.6; DB 10; Length 902;
Best Local Similarity 96.2%; Pred. No. 9e-105;
Matches 678; Conservative 0; Mismatches 19; Indels 8; Gaps 4;

QY      379 AGGATAATGCAGATTGATCACTTCTGTTTCACAGTGAGCTGATATACAGCGTTG 438
Db      44  AGGATAATGCAGATTGATCACTTCTGTTTCACAGTGAGCTGATATACAGCGTTG 103
QY      439 GATACGGTGGCTCAGTCCCTCATATTTGCTACATAGCTGGCCACCTAGAGGCTGCTG 498
Db      104 GATACGGTGGCTCAGTCCCTCATATTTGCTACATAGCTGGCCACCTAGAGGCTGCTG 163
QY      499 ATGTGCTGTTGCAACATGGAGCTAATGTCAATATTCAAGATGCAGTTTTTTTTCACATCCAT 558
Db      164 ATGTGCTGTTGCAACATGGAGCTAATGTCAATATTCAAGATGCAGTTTTTTTTCACATCCAT 223
QY      559 TGCATATTGCAGCGTACTATGGAATGAAACAGGTAATCGCCCTTCTTTTGAATTTGGTG 618
Db      224 TGCATATTGCAGCGTACTATGGAATGAAACAGGTAATCGCCCTTCTTTTGAATTTGGTG 283
QY      619 CTGATGTAATGTAAGTGGTGAAGTTGGAGATAGACCCCTCCACCTAGCATCTGCAAAAG 678
Db      284 CTGATGTAATGTAAGTGGTGAAGTTGGAGATAGACCCCTCCACCTAGCATCTGCAAAAG 343

```

```

QY      679 GATTTCTTGAATTTGCAAAACTCTTTGATGGAAGAAGGAGCAAGACAGATGTGAATGTC 738
Db      344 GATTTCTTGAATTTGCAAAACTCTTTGATGGAAGAAGGAGCAAGACAGATGTGAATGTC 403
QY      739 AAGATAATGAAGACCATGTCACACTCCATTTCTGTTCTCGATTTCGATTTGGACACCATGATATAG 798
Db      404 AAGATAATGAAGACCATGTCACACTCCATTTCTGTTCTCGATTTCGATTTGGACACCATGATATAG 463
QY      799 TTAAGTATCTGCTGCAAGAGTGAATTTGGAAGTTTCAACCTCATGTTCTTAAATATCTATGAG 858
Db      464 TTAAGTATCTGCTGCAAGAGTGAATTTGGAAGTTTCAACCTCATGTTCTTAAATATCTATGAG 523
QY      859 ATACCCCTTACACCTGGCATGCTACAATGGCAAAATTTGAAGTTCCCAAGAAATCATCC 918
Db      524 ATACCCCTTACACCTGGCATGCTACAATGGCAAAATTTGAAGTTCCCAAGAAATCATCC 582
QY      919 AATATACAGAACAGAAAGTCTGACTTAAGGAAACATCTTCAGTGAACACAGCTTTTCATA 978
Db      583 AATATACAGAACAGAAAGTCTGACTTAAGGAAACATCTTCAGTGAACACAGCTTTTCATA 641
QY      979 GTGTTGTACTATGGAAGAGCATGACCTAGTCAAAATTTCTTCTGATCAGAATGTCA 1038
Db      642 GTGTTGTACTATGGAAGAGCATGACCTAGTCAAAATTTCTTCTGATCAGAATGTCT 698
QY      1039 TAAACATCAACACCAAGGAGGATGGGCACACTGGATTACACT 1083
Db      699 TAGACATTCACACACAGGAA---GGATGGCACACTGGATTACTCT 740

RESULT 8
BX502807
LOCUS   DXFZp779C2070 r1 779 (synonym: hnccl) Homo sapiens cDNA clone
DEFINITION
mRNA sequence.
ACCESSION BX502807
VERSION   BX502807.1 GI:32024614
KEYWORDS EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 656)
AUTHORS   Bahr, A., Lauber, J., Mewes, H.W., Weil, B., Amid, C., Osanger, A.,
          Fobor, G., Han, M., and Wiemann, S.
          EST (Bahr, A., Lauber, J., Mewes, H.W., Weil, B., et al.)
          Unpublished (2003)
          Contact: MIPS
          MIPS
          Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
          This is the 5' sequence of the clone insert
          Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
          Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
          sequenced by Qiagen (Hilden/Germany) within the cDNA sequencing
          consortium of the German Genome Project.
          No sl sequence available.
          This clone (DXFZp779C2070) is available at the RZPD in Berlin.
          Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
          Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES             source
    source
        1..656
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="DKFZp779C2070"
            /tissue_type="liver"
            /dev_stage="fetal"
            /lab_host="DH10B"
            /note="Vector: pSport1_Sfi; Site_1: SfiIA; Site_2: SfiIB"

ORIGIN
Query Match      20.4%; Score 617.8; DB 13; Length 656;
Best Local Similarity 99.5%; Pred. No. 1.4e-104;

```

Matches	619;	Conservative	0;	Mismatches	3;	Indels	0;	Gaps	0;
QY	16	CGCGCCCTGGAGAAAGGAAGAACTTATAATAAATGGGAAATTTATAATCTAGACCAACCC	75						
Db	35	CTGCGCCCTGGAGAAAGGAAGAACTTATAATAAATGGGAAATTTATAATCTAGACCAACCC	94						
QY	76	AAACTTGTACTGATGAATGGAGAAAAGAGTCACTGAATCATATGTTATCACAATAGAAA	135						
Db	95	AAACTTGTACTGATGAATGGAGAAAAGAGTCACTGAATCATATGTTATCACAATAGAAA	154						
QY	136	GATTAGAAAGATGACCTCGACATCAAGGAAAAAGAACTGACAGAACTTAAGGAATATATTGTG	195						
Db	155	GATTAGAAAGATGACCTCGACATCAAGGAAAAAGAACTGACAGAACTTAAGGAATATATTGTG	214						
QY	196	GCTCTGATGAAGCCTTCAGTAAAGTCAATTTAAATTTACCGCACTGAAAATGGCGCTGCTC	255						
Db	215	GCTCTGATGAAGCCTTCAGTAAAGTCAATTTAAATTTACCGCACTGAAAATGGCGCTGCTC	274						
QY	256	TACTTCATTTATGTTGCATTTGTGGAGGCAAGAAATCACATATTCGAACCTCTTATGTTGA	315						
Db	275	TACTTCATTTATGTTGCATTTGTGGAGGCAAGAAATCACATATTCGAACCTCTTATGTTGA	334						
QY	316	AAGGCGCTCCGCCCATCTCGACTGACAGAAATCGATTTACAGCCTTGCATTTAGCAGTTT	375						
Db	335	AAGGCGCTCCGCCCATCTCGACTGACAGAAATCGATTTACAGCCTTGCATTTAGCAGTTT	394						
QY	376	ACAAGGATAATGACAGAAATGATCACTCTCTGCTTCACAGTGGAGCTGATATACAGAGG	435						
Db	395	ACAAGGATAATGACAGAAATGATCACTCTCTGCTTCACAGTGGAGCTGATATACAGAGG	454						
QY	436	TTGGATACGGTGGCGCTCACCTGCCCTCCATATTCCTACAAATAGCTGGCCACCTAGAGGCTG	495						
Db	455	TTGGATACGGTGGCGCTCACCTGCCCTCCATATTCCTACAAATAGCTGGCCACCTAGAGGCTG	514						
QY	496	CTGATGTGCTGTTGGAACATGGAGCTAAATGTCATATTTCAAGATGCAGTTTTCCTCCTC	555						
Db	515	CTGATGTGCTGTTGGAACATGGAGCTAAATGTCATATTTCAAGATGCAGTTTTCCTCCTC	574						
QY	556	CATTCCATATTGACAGCGTACTATGACATCAACAGGTAACTCGCCTCTCTTTTGAAATTTG	615						
Db	575	CATTCCATATTGACAGCGTACTATGACATCAACAGGTAACTCGCCTCTCTTTTGAAATTTG	634						
QY	616	GTGCTGATGTAATGTGAAGTGG	637						
Db	635	GTGCTGATGTAATGTGAAGTGG	656						

[illegible]

```

Fax: 216 361 9596
Email: scai@atersys.com
High quality sequence stop: 432.
Location/Qualifiers
1. .617
/organism="Homo sapiens"
/mol type="mRNA"
/db xref="taxon:9606"
/cell_line="HT1080"
/clone_lib="Athersys RAGE Library"
/note="See 'Creation of Genome-wide Protein Expression
Libraries using Random Activation of Gene Expression',
Nature Biotechnology, in press. Note that even though the
cell type indicated is HT1080, since a random activation
method was used, these sequence tags are not necessarily
expressed in HT1080 under normal circumstances."

```

## ORIGIN

Query Match	19.8%; Score 598.8; DB 12; Length 617;
Best Local Similarity	95.2%; Pred. No. 4.9e-101;
Matches 611; Conservative 0; Mismatches 4; Indels 1; Gaps 1;	
QY	2398 GTGCTGGACAAATATTCCCTCCAAGGTCTGTCTTTGGAGGAGATGAAGAAGTCTTCAAT 2457
DB	3 GTGCTGGACAAATATTCCTC-C-AAGTCTGTCTTTGGAGGAGATGAAGAAGTCTTCAAT 61
QY	2458 ACACACCATTGCACAAATATGGCTATGTATPCGATCCCATGAGCTCAATGCAATTTTCATT 2517
DB	62 ACACACCATTGCACAAATATGGCTATGTATPCGATCCCATGAGCTCAATGCAATTTTCATT 121
QY	2518 CTTGCCGAAATAGTAGCAGCTTTTGAGGACAGCAGCTGCACAGCATTCGGCGGTATACCTAAG 2577
DB	122 CTTGCCGAAATAGTAGCAGCTTTTGAGGACAGCAGCTGCACAGCATTCGGCGGTATACCTAAG 181
QY	2578 GAGAGTTTTTTTCCCCGAACTGCACACAACGATTCCAACCA CGGCAAGCTGGCTTCCAAC 2637
DB	182 GAGAGTTTTTTTCCCCGAACTGCACACAACGATTCCAACCA CGGCAAGCTGGCTTCCAAC 241
QY	2638 ATACATTTTACTCTCAAAGGTCTCCTTAATTTGGGCTGTGTTTTTACTGTGCCATTATAA 2697
DB	242 ATAACATTTTACTCTCAAAGGTCTCCTTAATTTGGGCTGTGTTTTTACTGTGCCATTATAA 301
QY	2698 TTCCCACATATTAGCAGGCTTTGGATTTGTGCCTAAGGAATAATATGCAAAAGAACCCAAG 2757
DB	302 TTCCCACATATTAGCAGGCTTTGGATTTGTGCCTAAGGAATAATATGCAAAAGAACCCAAG 361
QY	2758 ACAGAATGTATAGAAGATTTGTTTTTAATTTTGTAATTAATAAAAAAATTTAGATCGTT 2817
DB	362 ACAGAATGTATAGAAGATTTGTTTTTAATTTTGTAATTAATAAAAAAATTTAGATCGTT 421
QY	2818 ACTTGGAAATCGGACCTAAAGTCGTGGTGGACAGATAATAATATGTTTTCTCTGGGCTGA 2877
DB	422 ACTTGGAAATCGGACCTAAAGTCGTGGTGGACAGATAATAATATGTTTTCTCTGGGCTGA 481
QY	2878 ATTATGTAGACTTGTGTTTGACAGCTATGGGTTTATTTCTTTAGAACATTTGTTCTCT 2937
DB	482 ATTATGTAGACTTGTGTTTGACAGCTATGGGTTTATTTCTTTAGAACATTTGTTCTCT 541
QY	2938 TTTCTCATTAATGTTACTTCTTAGTGTGTCACCTCTGTGATTAAGAAATCTTTGGTGAATAG 2997
DB	542 TCTCTCATTAATGTTACTTCTTAGTGTGTCACCTCTGTGATTAAGAAATCTTTGGTGAATAG 601
QY	2998 AAAAAAAAAAAAAAAA 3013
DB	602 AAAAAAAAAAAAAAAA 617

RESULT 10	
BG216471	
LOCUS	
DEFINITION	
ACCESSION	
VERSION	

BG216471 617 bp mRNA linear EST 21-APR-2001  
 RST36050 Athersys RAGE Library Homo sapiens CDNA, mRNA sequence.  
 BG216471  
 BG216471.1 GI:13742380

```

KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
COMMENT
CONTACT: Scott J. Cain
3201 Carnegie Ave, Cleveland, OH 44115, USA
Tel: 216 431 9900
Fax: 216 361 9596
Email: scain@atherys.com
High quality sequence stop: 371.
FEATURES
source
1. 617
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/cell_line="HT1080"
/clone_lib="Athersys RAGE Library"
/notes="See 'Creation of Genome-wide Protein Expression
Libraries using Random Activation of Gene Expression',
Nature Biotechnology, in press. Note that even though the
cell type indicated is HT1080, since a random activation
method was used, these sequence tags are not necessarily
expressed in HT1080 under normal circumstances."
ORIGIN
Query Match 19.6%; Score 591.8; DB 12; Length 617;
Best Local Similarity 99.2%; Pred. No. 9.8e-100;
Matches 604; Conservative 0; Mismatches 4; Indels 1; Gaps 1;
QY 2402 TGGCAATATTCCTCTCAAGTCTGTCTTTGGAGGAGATGAAAGAAGTCTTCAATACAC 2461
DB 10 TGGCAATATTCCTCTC-TGGTCTGTCTTTGGAGGAGATGAAAGAAGTCTTCAATACAC 69
QY 2462 ACCATTGACAAATATGGCTATGTATCCGATCCCATGAGCTCAATGCATTTTCATTTG 2521
DB 69 ACCATTGACAAATATGGCTATGTATCCGATCCCATGAGCTCAATGCATTTTCATTTG 128
QY 2522 CCGAATATAGTACAGCTTTGAGGACAGCTGACAGCATTCGGCGTATACCTTAAGGAGA 2581
DB 129 CCGAATATAGTACAGCTTTGAGGACAGCTGACAGCATTCGGCGTATACCTTAAGGAGA 188
QY 2582 GTTTTTTCCCGAAGTACACGACGATTCACACGCGAAGCTGCTTCCAACTATATA 2641
DB 189 GTTTTTTCCCGAAGTACACGACGATTCACACGCGAAGCTGCTTCCAACTATATA 248
QY 2642 CATTTTACTCTCAAGGTCCTCTTAAATTTGGCTTTTACTTGTCTTATTTATTCCT 2701
DB 249 CATTTTACTCTCAAGGTCCTCTTAAATTTGGCTTTTACTTGTCTTATTTATTCCT 308
QY 2702 CCATTATAGCAGGCTTTGGATTTGTGCTTAAGGAATAATATCAAAAGAACCAAGACAG 2761
DB 309 CCATTATAGCAGGCTTTGGATTTGTGCTTAAGGAATAATATCAAAAGAACCAAGACAG 368
QY 2762 AATGATATAGGAATTTGTTTTTAAATTTTGTAAATTAATAAATAATTTAGATCGTTACTT 2821
DB 369 AATGATATAGGAATTTGTTTTTAAATTTTGTAAATTAATAAATAATTTAGATCGTTACTT 428
QY 2822 GGAATGAGGCTTAAGTCTGTGTGGACAGATAATAATATGTTTTCTGGGCTGAATTA 2881
DB 429 GGAATGAGGCTTAAGTCTGTGTGGACAGATAATAATATGTTTTCTGGGCTGAATTA 488
QY 2882 TGAGACTTGTGTTTACAGCTATGGTGTATTTCTTAGAACATTTCTTTTTC 2941
DB 489 TGAGACTTGTGTTTACAGCTATGGTGTATTTCTTAGAACATTTCTTTTTC 548
QY 2942 TCATTATGTTACTTCTAGTGTTCACCTCTGTGATTAAGATCTTTTGGTGAATAGAAA 3001
DB 549 TCATTATGTTACTTCTAGTGTTCACCTCTGTGATTAAGATCTTTTGGTGAATAGAAA 608
QY 3002 AAAAAAAA 3010
DB 609 AAAAAAAA 617
RESULT 11
BG215938 785 bp mRNA linear EST 21-APR-2001
LOCUS RGT35624 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.
DEFINITION
ACCESSION BG215938
VERSION BG215938.1 GI:13741959
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Mammalia; Euthera; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
AUTHORS Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R.,
Cain,S., Leventhal,C., Thornton,M., Ramachandran,R.,
Whittington,J., Lerner,L., Costanzo,D., McElligott,K., Booser,S.,
Mays,R., Smith,E., Veloso,N., Klika,A., Hess,J., Cothren,K., Lo,K.,
Offenbacher,J., Danzig,J. and Ducar,M.
TITLE Creation of genome-wide protein expression libraries using random
activation of gene expression
JOURNAL Nat. Biotechnol. 19 (5), 440-445 (2001)
MEDLINE 21227151
PUBMED 11329013
COMMENT Contact: Scott J. Cain
Athersys, Inc.
3201 Carnegie Ave, Cleveland, OH 44115, USA
Tel: 216 431 9900
Fax: 216 361 9596
Email: scain@atherys.com
High quality sequence stop: 550.
FEATURES
source
1. 785
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/cell_line="HT1080"
/clone_lib="Athersys RAGE Library"
/notes="See 'Creation of Genome-wide Protein Expression
Libraries using Random Activation of Gene Expression',
Nature Biotechnology, in press. Note that even though the
cell type indicated is HT1080, since a random activation
method was used, these sequence tags are not necessarily
expressed in HT1080 under normal circumstances."
ORIGIN
Query Match 19.5%; Score 590.8; DB 12; Length 785;
Best Local Similarity 99.7%; Pred. No. 1.4e-99;
Matches 592; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 2420 AGGTCTGTCTTTGGAGGAGATGAAAGAAGTCTTCAATACACACCCATGCAATATGG 2479
DB 23 AGGTCTGTCTTTGGAGGAGATGAAAGAAGTCTTCAATACACACCCATGCAATATGG 82
QY 2480 CTATGATCCGATCCCATGAGCTCAATGCAATTTTCATTTCTTCCGAAATAGTAGAGCTT 2539
DB 83 CTATGATCCGATCCCATGAGCTCAATGCAATTTTCATTTCTTCCGAAATAGTAGAGCTT 142
QY 2540 TGAGGACAGCAGCTGACAGCATTCGGCGTATACCTTAAGGAGATTTTTCCTCCGACTGA 2599
DB 143 TGAGGACAGCAGCTGACAGCATTCGGCGTATACCTTAAGGAGATTTTTCCTCCGACTGA 202

```





Db 530 CTGATGTCCTGTTGCAACATGGAGCTAATGCAATATTCAGATGCAGTGTTCCTCACTC 589  
QY 556 CATTGCATATTGACGCTACTATGACATGACAGGTAACTCGCCTCT 604  
Db 590 CATTGCATATTGACGCTACTATGACATGACAGGTAACTCGCCTCT 638

RESULT 15  
BG181490  
LOCUS BG181490 785 bp mRNA linear EST 21-APR-2001  
DEFINITION R31338 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.  
ACCESSION BG181490  
VERSION BG181490.1 GI:13703177  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 785)  
Harrington, J.J., Sherf, B., Rundlett, S., Jackson, P.D., Perry, R.,  
Cain, S., Leventhal, C., Thornton, M., Ramachandran, R.,  
Whittington, J., Lerner, L., Costanzo, D., McElligott, K., Booser, S.,  
Mays, R., Smith, E., Veloso, N., Klika, A., Hess, J., Cothren, K., Lo, K.,  
Offenbacher, J., Danzig, J. and Ducar, M.,  
Creation of genome-wide protein expression libraries using random  
activation of gene expression  
Nat. Biotechnol. 19 (5), 440-445 (2001)  
21227151  
11329013  
Contact: Scott J. Cain  
Athersys, Inc.  
3201 Carnegie Ave, Cleveland, OH 44115, USA  
Tel: 216 431 9900  
Fax: 216 361 9596  
Email: scain@atersys.com  
High quality sequence stop: 481.

FEATURES  
source  
1..785  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/cell\_line="HT1080"  
/clone\_lib="Athersys RAGE Library"  
/notes="See 'Creation of Genome-wide Protein Expression  
Libraries using Random Activation of Gene Expression',  
Nature Biotechnology, in press. Note that even though the  
cell type indicated is HT1080, since a random activation  
method was used, these sequence tags are not necessarily  
expressed in HT1080 under normal circumstances."

ORIGIN  
Query Match 19.4%; Score 587.2; DB 12; Length 785;  
Best Local Similarity 97.1%; Pred. No. 6.4e-99;  
Matches 598; Conservative 0; Mismatches 18; Indels 0; Gaps 0;  
QY 2398 GTGCTGGACAATATTCCTCTCAAGGCTCTCTCTTTGGAGGAGATGAAAGAGTCTTCAAT 2457  
Db 4 GTGCTGGACAATATTCCTCTTAAGGTCTCTCTTTGGAGGAGATGAAAGAGTCTTCAAT 63  
QY 2458 ACACACCATTTGACAAATATGGCTATGATCCGATCCCATGAGCTCAATGCAATTTTCATT 2517  
Db 64 ACACACCATTTGACAAATATGGCTATGATCCGATCCCATGAGCTCAATGCAATTTTCATT 123  
QY 2518 CTTGCCGAATAGTAGCAGCTTTGAGGACAGAGCTGACAGCATTCGGCGTATACCTTAAG 2577  
Db 124 CTTGCCGAATAGTAGCAGCTTTGAGGACAGAGCTGACAGCATTCGGCGTATACCTTAAG 183  
QY 2578 GAGAGTTTTTCCCGAATGACAGCAAGCATTCACACCGGCAAGCTGGCTTCCAACT 2637  
Db 184 GAGAGTTTTTCCCGAATGACAGCAAGCATTCACACCGGCAAGCTGGCTTCCAACT 243  
QY 2638 ATACATTTTACTCTCAAGGCTCTCTTAATTTGGCTGTGTTTACTGTTCCTATTAA 2697

Search completed: September 6, 2004, 07:38:00  
Job time : 5038.83 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 5, 2004, 19:27:21 ; Search time 631.459 Seconds  
(without alignments)  
16852.608 Million cell updates/sec

Title: US-10-626-173-3  
Perfect score: 2505  
Sequence: 1 atgggaataataaatttag.....gcagctttgagacagcgc 2505

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N Geneseq\_29Jan04:\*  
1: Geneseqn1980s:\*  
2: Geneseqn1990s:\*  
3: Geneseqn2000s:\*  
4: Geneseqn2001as:\*  
5: Geneseqn2001bs:\*  
6: Geneseqn2002s:\*  
7: Geneseqn2003as:\*  
8: Geneseqn2003bs:\*  
9: Geneseqn2003cs:\*  
10: Geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2505	100.0	2505	3	AAA47607	Aaa47607 Human CAR
2	2505	100.0	2508	4	Aaf44702	Aaf44702 Novel pro
3	2505	100.0	3025	3	AAA47606	Aaa47606 Human CAR
4	2505	100.0	3025	7	ACC48580	Acc48580 Human car
5	2455.4	98.0	2788	5	AAS80720	Aas80720 DNA encod
6	1892.2	75.5	2505	3	AAA47609	Aaa47609 Rat CAR
7	1892.2	75.5	3026	3	AAA47608	Aaa47608 Rat CAR
8	1892.2	75.5	3026	7	ACC48581	Acc48581 Rat cardi
9	1772	70.7	2024	8	AAD57334	Aad57334 Human kin
10	466.4	18.6	668	6	ABT09088	Abt09088 Phase-1 R
11	458.4	18.3	633	4	AAS33234	Aas33234 DNA encod
12	432	17.2	640	4	AAS33357	Aas33357 DNA encod
13	135.6	5.4	1557	5	AAS80722	Aas80722 DNA encod
14	106.4	4.2	2283	3	ACC48526	Aac48526 Arabidops
15	100.2	4.0	1704	7	ADA69514	Ada69514 Rice gene
16	98	3.9	1662	6	ABZ13209	Abz13209 Arabidops
17	91.4	3.6	3323	9	ADB62159	Adb62159 Human cdn
18	87.8	3.5	1845	7	ADA70011	Ada70011 Rice gene
19	87.4	3.5	1299	4	AAS33288	Aas33288 DNA encod
20	87.4	3.5	1299	5	ABK72043	Abk72043 Human cdn
21	87.4	3.5	1299	6	ABAI4324	Abai4324 Human ner
22	87.4	3.5	1299	6	ABK91635	Abk91635 CDNA enco
23	85.6	3.4	3553	9	ADC39121	Adc39121 Novel hum

24	85.6	3.4	4273	4	ABS54855	AbS54855 Human ank
25	85.6	3.4	4519	3	AAD51564	Aad51564 Human str
26	84.6	3.4	1514	3	AAC47638	Aac47638 Arabidops
27	84.2	3.4	5175	6	AAX99407	Aax99407 DNA of AP
28	84	3.4	2915	6	ABN59866	Abn59866 Novel hum
29	84	3.4	13329	5	AAS71562	Aas71562 DNA encod
30	84	3.4	14770	6	ABJ61893	Abj61893 Colon ade
31	84	3.4	14770	6	ABJ68857	Abj68857 Kidney ca
32	84	3.4	14770	6	ABJ68858	Abj68858 Kidney ca
33	84	3.4	14770	6	ADB31314	Adb31314 Testoster
34	84	3.4	14797	5	AAS65159	Aas65159 DNA encod
35	84	3.4	15187	5	AAS86131	Aas86131 DNA encod
36	82.6	3.3	6278	4	AAK52293	Aak52293 Human pcl
37	80.8	3.2	1797	5	AAK590586	Aak590586 Degenerat
38	80.8	3.2	1797	6	AAK36756	Aak36756 Human rin
39	80.6	3.2	3160	4	AAC90485	Aac90485 Rice EDR1
40	80.2	3.2	883	6	ABK65288	Abk65288 Arabidops
41	80.2	3.2	883	9	ADD30938	Add30938 Plant yle
42	80.2	3.2	883	9	ADG37244	Adg37244 Plant yle
43	79.6	3.2	3438	9	ADB62405	Adb62405 Human cdn
44	79.4	3.2	1914	7	ADA69560	Ada69560 Rice gene
45	79.4	3.2	2187	7	ADA70219	Ada70219 Rice gene

ALIGNMENTS

RESULT 1  
AAA47607  
ID AAA47607 standard; cDNA; 2505 BP.  
XX  
AC AAA47607;  
XX  
DT 20-OCT-2000 (first entry)  
XX  
DE Human CAR (Cardiac related Ankyrin-Repeat Protein Kinase) cDNA.

KW Cardiac related ankyrin repeat protein kinase; CAR; cytoskeleton;  
KW cardiac cell growth factor receptor; cell differentiation; modulator;  
KW regulator; detection; cellular proliferation; cardiovascular disorder;  
KW heart failure; hypertension; cancer; sarcoma; ds.  
XX Homo sapiens.  
XX Key Location/Qualifiers  
FH 1. .2505  
CDS /\*tag= a  
FT /product= "Human CAR"

PN WO200034330-A1.

XX 15-JUN-2000.

XX 10-DEC-1999; 99WO-US029465.

PR 11-DEC-1998; 98US-0111938P.

PR 14-APR-1999; 99US-00291839.

XX (MILL-) MILLENNIUM PHARM INC.

XX Raju J;

DR WFI; 2000-431275/37.

DR P-PSDB; AAB01470.

XX New polynucleotide encoding cardiac-related ankyrin-repeat protein

PT kinase, useful for treating disorders such as cardiovascular disorders,

PT e.g. heart failure and cell differentiation disorders, e.g. cancer.

XX Claim 1; Page 143-147; 16ipp; English.

XX CARK polypeptides are regulators of signal transmission from cellular

CC receptors (e.g. cardiac cell growth factor receptor) and cytoskeleton

function. They also act as modulators of cellular differentiation and cell death. Nucleic acids, or their fragments encoding CARP polypeptides are useful for detecting CARP nucleic acids especially mRNA, in a sample. CARP is useful for treating disorders associated with upregulation or downregulation of cellular proliferation such as, cardiovascular disorders (heart failure and hypertension) and disorders associated with cell differentiation such as cancer and sarcoma

Sequence 2505 BP; 721 A; 532 C; 554 G; 698 T; 0 U; 0 Other;

```
Query Match      100.0%; Score 2505; DB 3; Length 2505;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy	1	ATGGGAAATTATAAATCTAGACCAACCAAACTTGTACTGATGAATGGAAGAAAAGTC	60
Db	1	ATGGGAAATTATAAATCTAGACCAACCAAACTTGTACTGATGAATGGAAGAAAAGTC	60
Qy	61	AGTGAATCATATGTTTATCACAATAGAAGAATTAGAAGATGACCTGCAGATCAAGGAAAA	120
Db	61	AGTGAATCATATGTTTATCACAATAGAAGAATTAGAAGATGACCTGCAGATCAAGGAAAA	120
Qy	121	GAACTGCAGAAACTAAGGAATATATTTGGCTCTGATGAAGCCTTCAGTAAAGTCAATTTA	180
Db	121	GAACTGCAGAAACTAAGGAATATATTTGGCTCTGATGAAGCCTTCAGTAAAGTCAATTTA	180
Qy	181	AATTACCGCACGAAATGGGCTGCTCTACTTCATTTATGTTGGCATTTGTGGAGGCAAG	240
Db	181	AATTACCGCACGAAATGGGCTGCTCTACTTCATTTATGTTGGCATTTGTGGAGGCAAG	240
Qy	241	AAATCACATATTCGAACCTTTATGTTGAAAGGGCTCGGCCATCTCGACTGACAAGAAAT	300
Db	241	AAATCACATATTCGAACCTTTATGTTGAAAGGGCTCGGCCATCTCGACTGACAAGAAAT	300
Qy	301	GGATTTACAGCCTTCGCAATTTAGCAGTTTACAGGATAATGCAGAAATGATCACTTCTCTG	360
Db	301	GGATTTACAGCCTTCGCAATTTAGCAGTTTACAGGATAATGCAGAAATGATCACTTCTCTG	360
Qy	361	CTTCACAGTGGAGCTGATATACAGCAGGTTGGATACGGTGGCCTCACTGCCCTCCATATT	420
Db	361	CTTCACAGTGGAGCTGATATACAGCAGGTTGGATACGGTGGCCTCACTGCCCTCCATATT	420
Qy	421	GCTCAATAGCTGGCCACCTAGAGCTGCTGATGCTGTTGCAAATCGAGCTAATGTC	480
Db	421	GCTCAATAGCTGGCCACCTAGAGCTGCTGATGCTGTTGCAAATCGAGCTAATGTC	480
Qy	481	AATATTCAGATGCAAGTCTTTTTTCACTCCATTGCATATTCACGCTACTATGGACATGAA	540
Db	481	AATATTCAGATGCAAGTCTTTTTTCACTCCATTGCATATTCACGCTACTATGGACATGAA	540
Qy	541	CAGGTAACCTCGCTCTTTTGAAATTTGGTCTGATGTAATAATGTAAGTGGTGAAGTTGA	600
Db	541	CAGGTAACCTCGCTCTTTTGAAATTTGGTCTGATGTAATAATGTAAGTGGTGAAGTTGA	600
Qy	601	GATAGACCCCTCCACTAGCATCTGCBAAGGATTCCTTGAATATTCGAAAACCTCTTGATG	660
Db	601	GATAGACCCCTCCACTAGCATCTGCBAAGGATTCCTTGAATATTCGAAAACCTCTTGATG	660
Qy	661	GAAGAAGGCAGCAAAAGCAGATGTAATGCTCAAGATAATGAAGACCATGTCCCACTCCAT	720
Db	661	GAAGAAGGCAGCAAAAGCAGATGTAATGCTCAAGATAATGAAGACCATGTCCCACTCCAT	720
Qy	721	TTCTGTTCTCGATTTGGACACCATGATATAGTTAAGTATCTGCTGCAAGATGATTTGGA	780
Db	721	TTCTGTTCTCGATTTGGACACCATGATATAGTTAAGTATCTGCTGCAAGATGATTTGGA	780
Qy	781	GTTCAACCTCATGTTGTTAAATATCTATGGAGATACCCCTTACACCTGGCAGCTGCTACAT	840
Db	781	GTTCAACCTCATGTTGTTAAATATCTATGGAGATACCCCTTACACCTGGCAGCTGCTACAT	840
Qy	841	GGCAAAATTTGAAGTTGCCAAGGAATCATCAAAATATCAGGAACAGAAAGCTGCTAAG	900
Db	841	GGCAAAATTTGAAGTTGCCAAGGAATCATCAAAATATCAGGAACAGAAAGCTGCTAAG	900



```
Db 661 ||||| GAAGAGGCGCAAGCAGATGTGAATGCTCAGATAATGAACACATGTCACCATCCAT 720
Qy 721 TTCTGTTCTCGATTGGACACCATGATATAGTTAAAGTATCTGTCGCAAAAGTATTGGAA 780
Db 721 TTCTGTTCTCGATTGGACACCATGATATAGTTAAAGTATCTGTCGCAAAAGTATTGGAA 780
Qy 781 GTTCAACCTCATGTTGTTAATATCTATGAGAGATACCCCTTACACCTGGCATGCTACAT 840
Db 781 GTTCAACCTCATGTTGTTAATATCTATGAGAGATACCCCTTACACCTGGCATGCTACAT 840
Qy 841 GGCAGAAATTTGAAGTTGCCAAGGAAATCATCCAAATATCAGGAACAGAAAGTCTGACTAAG 900
Db 841 GGCAGAAATTTGAAGTTGCCAAGGAAATCATCCAAATATCAGGAACAGAAAGTCTGACTAAG 900
Qy 901 GAAACATCTTTCAGTGAAACAGCTTTTCATAGTGTCTGTGATACCTATGCGAAGACATGAC 960
Db 901 GAAACATCTTTCAGTGAAACAGCTTTTCATAGTGTCTGTGATACCTATGCGAAGACATGAC 960
Qy 961 CTAGTCAAAATTTCTTTGATCAGAAATGTCATAAATCAACCAACCAAGGAGGATGG 1020
Db 961 CTAGTCAAAATTTCTTTGATCAGAAATGTCATAAATCAACCAACCAAGGAGGATGG 1020
Qy 1021 CACACTGGATTACACTCTGCTTGTCTACCCAGGTCACATTCGCTGTTTCACTTCTTACTG 1080
Db 1021 CACACTGGATTACACTCTGCTTGTCTACCCAGGTCACATTCGCTGTTTCACTTCTTACTG 1080
Qy 1081 GATATGGAGCTGATATCAATCTAGTGGCTTGTATCCAGCAGAGTCTAGTGGTGAATA 1140
Db 1081 GATATGGAGCTGATATCAATCTAGTGGCTTGTATCCAGCAGAGTCTAGTGGTGAATA 1140
Qy 1141 GATGAGCAGACATGTTTGTGATGCTGTTATGAAAAAGGCGATGATGCCATTTGTCACATC 1200
Db 1141 GATGAGCAGACATGTTTGTGATGCTGTTATGAAAAAGGCGATGATGCCATTTGTCACATC 1200
Qy 1201 CTGAAGCATTTAAGACACCAAGATGAATTTGCCCTGTAATGAATATTTCTAGCCTGGA 1260
Db 1201 CTGAAGCATTTAAGACACCAAGATGAATTTGCCCTGTAATGAATATTTCTAGCCTGGA 1260
Qy 1261 GGAGATGGCTCTATGTTGCTTCCATCACCTTGGGGAAGATTAAAAAGCATGACAAA 1320
Db 1261 GGAGATGGCTCTATGTTGCTTCCATCACCTTGGGGAAGATTAAAAAGCATGACAAA 1320
Qy 1321 GAGAAGCAGATATTTCTCCTTAAGAGCTGGAATGCTTTCATATTTCCATTTTCAGCTC 1380
Db 1321 GAGAAGCAGATATTTCTCCTTAAGAGCTGGAATGCTTTCATATTTCCATTTTCAGCTC 1380
Qy 1381 TCAGAAATGAGTTCCATGAGATTTATGGCTCAGGTTCTTTTGGGAAGTATATAAGGA 1440
Db 1381 TCAGAAATGAGTTCCATGAGATTTATGGCTCAGGTTCTTTTGGGAAGTATATAAGGA 1440
Qy 1441 CGATGCGAAATAAAATAGTGGCTATAAAACGTTATCGAGCCAAATCTTACTGCTCCAAG 1500
Db 1441 CGATGCGAAATAAAATAGTGGCTATAAAACGTTATCGAGCCAAATCTTACTGCTCCAAG 1500
Qy 1501 TCAGATGTGATATGTTTTCGCGAGAGGTGTCATTTCTGCGCAGCTCAATCATCCCTGC 1560
Db 1501 TCAGATGTGATATGTTTTCGCGAGAGGTGTCATTTCTGCGCAGCTCAATCATCCCTGC 1560
Qy 1561 GTAATTCAGTTTGTGGGTGCTTGTGTAATGATCCAGCAGGTTTGGCAATTTGCTACTCAA 1620
Db 1561 GTAATTCAGTTTGTGGGTGCTTGTGTAATGATCCAGCAGGTTTGGCAATTTGCTACTCAA 1620
Qy 1621 TACATATCAGGGGTTCTGTTCTCCCTTCATGAGCAGAGAGGATTTCTGATTTG 1680
Db 1621 TACATATCAGGGGTTCTGTTCTCCCTTCATGAGCAGAGAGGATTTCTGATTTG 1680
Qy 1681 CAGTCTAAATTAATTTAGTGTAGATGTTGCCAAAGGATGGAGTACCTTCCACACCTG 1740
Db 1681 CAGTCTAAATTAATTTAGTGTAGATGTTGCCAAAGGATGGAGTACCTTCCACACCTG 1740
Qy 1741 ACACAGCCAATTAATACATCTGTGACTTGAACAGTCAAAATTTCTCTATGAGGATGG 1800
```

```
Db 1741 ACACAGCCAATTAATACATCTGTGACTTGAACAGTCAAAATTTCTCTATGAGGATGG 1800
Qy 1801 CATGCTGTGTGCGAGATTTTGGAGAAATCAAGATTTTCAAGTCTCTGATGAAGACAAC 1860
Db 1801 CATGCTGTGTGCGAGATTTTGGAGAAATCAAGATTTTCAAGTCTCTGATGAAGACAAC 1860
Qy 1861 ATGACAAAACAACTGGACCTCCGTTGGATGGCTCTGAGGTTGTCACGAGTCACT 1920
Db 1861 ATGACAAAACAACTGGACCTCCGTTGGATGGCTCTGAGGTTGTCACGAGTCACT 1920
Qy 1921 CGGTACACCATCAAGACAGATGTCTTTCAGCTATGCTGTGTCTGTGGGAAATTTCTCACT 1980
Db 1921 CGGTACACCATCAAGACAGATGTCTTTCAGCTATGCTGTGTCTGTGGGAAATTTCTCACT 1980
Qy 1981 GGGGAAATTTCCATTTCCCTCATCTCAAGCCAGCGGCTGGGAGCAGACATGGCTTACCAC 2040
Db 1981 GGGGAAATTTCCATTTCCCTCATCTCAAGCCAGCGGCTGGGAGCAGACATGGCTTACCAC 2040
Qy 2041 CACATCAGACCTCCCATTTGGCTATTCCATTCCCAAGCCCATATCATCTCTGCTGATACGA 2100
Db 2041 CACATCAGACCTCCCATTTGGCTATTCCATTCCCAAGCCCATATCATCTCTGCTGATACGA 2100
Qy 2101 GGGTGAACGCGATGCTCTGAAAGAAACCCGAAATTTCTGAAGTTGTCTGAAGTTAGAA 2160
Db 2101 GGGTGAACGCGATGCTCTGAAAGAAACCCGAAATTTCTGAAGTTGTCTGAAGTTAGAA 2160
Qy 2161 GAGTGTCTCTGCAACATTCAGCTGATGCTCTCTGATCAAGTAAACAGCAGTGGGTCTCTC 2220
Db 2161 GAGTGTCTCTGCAACATTCAGCTGATGCTCTCTGATCAAGTAAACAGCAGTGGGTCTCTC 2220
Qy 2221 TCACCTTCTTCTTCTTCTGATTTCCCTTCCCTCAAGGTCTGCTTTTGGAGGATGAAAGA 2400
Db 2221 TCACCTTCTTCTTCTTCTGATTTCCCTTCCCTCAAGGTCTGCTTTTGGAGGATGAAAGA 2400
Qy 2281 GCAGCATTAAGAGTGGTTTCCGAATTTGGAATGCTCTTAATGCAAGTCTCTATGCTGCT 2340
Db 2281 GCAGCATTAAGAGTGGTTTCCGAATTTGGAATGCTCTTAATGCAAGTCTCTATGCTGCT 2340
Qy 2341 TTCTCCCAAGTGTGCGACAATATTCTCTCAAGGTCTGCTTTTGGAGGATGAAAGA 2400
Db 2341 TTCTCCCAAGTGTGCGACAATATTCTCTCAAGGTCTGCTTTTGGAGGATGAAAGA 2400
Qy 2401 AGTCTTCAATACACACCCATTCACAAATATGCTGATGATCGATCCCATGAGCTCAATG 2460
Db 2401 AGTCTTCAATACACACCCATTCACAAATATGCTGATGATCGATCCCATGAGCTCAATG 2460
Qy 2461 CATTTTTCATTTTGGCGAAATAGTAGCAGCTTTGAGGACAGCAGC 2505
Db 2461 CATTTTTCATTTTGGCGAAATAGTAGCAGCTTTGAGGACAGCAGC 2505
```

## RESULT 3

AAA47606  
ID AAA47606 standard; cDNA; 3025 bp.

XX AC AAA47606;

XX DT 20-OCT-2000 (first entry)

XX DE Human CARK (Cardiac related Ankyrin-Repeat Protein Kinase) cDNA.

XX KW Cardiac related ankyrin repeat protein kinase; CARK; cytoskeleton;  
KW cardiac cell growth factor receptor; cell differentiation; modulator;  
KW regulator; detection; cellular proliferation; cardiovascular disorder;  
XX heart failure; hypertension; cancer; sarcoma; ds.

OS Homo sapiens.

XX PH Location/Qualifiers

XX FT 48..2555

XX FT /tag= a

XX FT /product= "Human CARK"

PN WO2000034330-A1.  
XX 15-JUN-2000. 99WO-US029465.  
XX 10-DEC-1999; 99WO-US029465.  
XX 11-DEC-1999; 98US-0111938P.  
PR 14-APR-1999; 99US-00291839.  
XX (MILL-) MILLENNIUM PHARM INC.  
XX Raju J;  
XX WPI; 2000-431275/37.  
DR P-PSDB; AAB01470.  
XX New polynucleotide encoding cardiac-related ankyrin-repeat protein  
PT kinase, useful for treating disorders such as cardiovascular disorders,  
PT e.g. heart failure and cell differentiation disorders, e.g. cancer.  
XX Claim 1; Fig 1; 161pp; English.  
XX CARP polypeptides are regulators of signal transmission from cellular  
CC receptors (e.g. cardiac cell growth factor receptor) and cytoskeleton  
CC function. They also act as modulators of cellular differentiation and  
CC cell death. Nucleic acids, or their fragments encoding CARP polypeptides  
CC are useful for detecting CARP nucleic acids especially mRNA, in a sample.  
CC CARP is useful for treating disorders associated with upregulation or  
CC downregulation of cellular proliferation such as, cardiovascular  
CC disorders (heart failure and hypertension) and disorders associated with  
CC cell differentiation such as cancer and sarcoma  
XX Sequence 3025 BP; 882 A; 622 C; 655 G; 866 T; 0 U; 0 Other;  
Query Match 100.0%; Score 2505; DB 3; Length 3025;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGGGAATTATTAATCTAGACCAACCCAACTTGTACTGATGAATGGAAGAAAAGTC 60  
DB 48 ATGGGAAATTATAAATCTAGACCAACCCAACTTGTACTGATGAATGGAAGAAAAGTC 107  
QY 61 AGTGAATCATATGTTATCAATAGAAAGATTAGAGATGACCTCAGATCAAGGAAAAA 120  
DB 108 AGTGAATCATATGTTATCAATAGAAAGATTAGAGATGACCTCAGATCAAGGAAAAA 167  
QY 121 GAATGACAGAACTAAGGAATATATTGGCTCTGATGAAGCCCTTCAGTAAAGTCATTTA 180  
DB 168 GAATGACAGAACTAAGGAATATATTGGCTCTGATGAAGCCCTTCAGTAAAGTCATTTA 227  
QY 181 AATTACCGCACTGAAATGGGCTGTCTCTACTTTCATTTATGTTGATTTGGAGGCAAG 240  
DB 228 AATTACCGCACTGAAATGGGCTGTCTCTACTTTCATTTATGTTGATTTGGAGGCAAG 287  
QY 241 AATATCATATTCGAATCTTATGTTGAAAGGGCTCCGCCCCATCTCGACTGACAGAAAT 300  
DB 288 AATATCATATTCGAATCTTATGTTGAAAGGGCTCCGCCCCATCTCGACTGACAGAAAT 347  
QY 301 GGATTTACAGCCCTTGCATTTAGCAGTTTACAGGATATGCGAATTTGATCACTTCTCTG 360  
DB 348 GGATTTACAGCCCTTGCATTTAGCAGTTTACAGGATATGCGAATTTGATCACTTCTCTG 407  
QY 361 CTTTCACAGTGGAGCTGATATACAGCAGGTTGGATACGGTGGCTCAGCTCCCTCCATATT 420  
DB 408 CTTTCACAGTGGAGCTGATATACAGCAGGTTGGATACGGTGGCTCAGCTCCCTCCATATT 467  
QY 421 GCTACAAATAGTCGCCACCTAGAGGCTGTGATGTGCTGTGCTTGCACATGGAGCTTAATGC 480  
DB 468 GCTACAAATAGTCGCCACCTAGAGGCTGTGATGTGCTGTGCTTGCACATGGAGCTTAATGC 527  
QY 481 AATATTCAAGATCAGTCTTTTCTACTCCATTTGCATTTGACGCTTACTATGGACATGAA 540  
DB 528 AATATTCAAGATCAGTCTTTTCTACTCCATTTGCATTTGACGCTTACTATGGACATGAA 587

QY 541 CAGGTAACTCGCTTCTTTTGAATTTGGTGTGATGTAATGTAAGTGGTGAAGTTGGA 600  
DB 588 CAGTAACTCGCTTCTTTTGAATTTGGTGTGATGTAATGTAAGTGGTGAAGTTGGA 647  
QY 601 GATAGACCCCTCCACTAGCATCTGCAAAAGGATTCCTTGAATATTTGCAAACTCTTGATG 660  
DB 648 GATAGACCCCTCCACTAGCATCTGCAAAAGGATTCCTTGAATATTTGCAAACTCTTGATG 707  
QY 661 GAAGAAGGCAGCAAGAGCAGATGTGAATGCTCAAGATAATGAAGACCATGTCCCACTCCAT 720  
DB 708 GAAGAAGGCAGCAAGAGCAGATGTGAATGCTCAAGATAATGAAGACCATGTCCCACTCCAT 767  
QY 721 TTCTGTTCTCGATTTGGACACCATGATATAGTAAATATCTGCTGCAAAAGTGAATTTGAA 780  
DB 768 TTCTGTTCTCGATTTGGACACCATGATATAGTAAATATCTGCTGCAAAAGTGAATTTGAA 827  
QY 781 GTTCAACCTCATGTTGTTAAATATCTATGGAGATACCCCTTACACCTTGGCATGCTACAA 840  
DB 828 GTTCAACCTCATGTTGTTAAATATCTATGGAGATACCCCTTACACCTTGGCATGCTACAA 887  
QY 841 GGCATAATTTGAAAGTTGCCAAGGAAATCATCCAAATATCAGGAACAGAAAGTCTGACTAAG 900  
DB 888 GGCATAATTTGAAAGTTGCCAAGGAAATCATCCAAATATCAGGAACAGAAAGTCTGACTAAG 947  
QY 901 GAAACATCTTCAGTGAACACAGCTTTTCATAGTGTCTTACTATGTCGCAAGCATTTGAC 960  
DB 948 GAAACATCTTCAGTGAACACAGCTTTTCATAGTGTCTTACTATGTCGCAAGCATTTGAC 1007  
QY 961 CTAGTCAAAATTTCTTCTGATCAGAAATGTCTAAACATCAACCAACCAAGGAGGATGGG 1020  
DB 1008 CTAGTCAAAATTTCTTCTGATCAGAAATGTCTAAACATCAACCAACCAAGGAGGATGGG 1067  
QY 1021 CACACTGGATTAACATCTGCTTGTACCAAGTGTACATTCGCTGCTGTTTCACTTCTTACTG 1080  
DB 1068 CACACTGGATTAACATCTGCTTGTACCAAGTGTACATTCGCTGCTGTTTCACTTCTTACTG 1127  
QY 1081 GATAATGGAGCTGATATGAATCTAGTGTGTTGATCCAGCAGCTCTAGTGGTGAAGAAA 1140  
DB 1128 GATAATGGAGCTGATATGAATCTAGTGTGTTGATCCAGCAGCTCTAGTGGTGAAGAAA 1187  
QY 1141 GATGACGACATGTTTGTGATGTTGGCTTTATGAAAAGGGCATGATGCCATTTGTCACTC 1200  
DB 1188 GATGACGACATGTTTGTGATGTTGGCTTTATGAAAAGGGCATGATGCCATTTGTCACTC 1247  
QY 1201 CTGAGCATTTAAGAGACCAAGATGAAATGCTGCTGATGATGATTTCTTCAGCTCGGA 1260  
DB 1248 CTGAGCATTTAAGAGACCAAGATGAAATGCTGCTGATGATGATTTCTTCAGCTCGGA 1307  
QY 1261 GGAGATGGCTCTCTATGTTCTTCCATCACCTTTGGGAAAGATTAAAGCATGACAAAA 1320  
DB 1308 GGAGATGGCTCTCTATGTTCTTCCATCACCTTTGGGAAAGATTAAAGCATGACAAAA 1367  
QY 1321 GAGAAGGCAGATATTCTCTCTCTAAGAGCTGGATTGCTTCCATTTCCATTTCCAGCTC 1380  
DB 1368 GAGAAGGCAGATATTCTCTCTCTAAGAGCTGGATTGCTTCCATTTCCATTTCCAGCTC 1427  
QY 1381 TCAGAAATTTGAGTTCCATGAGATTTTGGCTCAGCTTCTTTTGGGAAAGTATATAAGGA 1440  
DB 1428 TCAGAAATTTGAGTTCCATGAGATTTTGGCTCAGCTTCTTTTGGGAAAGTATATAAGGA 1487  
QY 1441 CGATGCAAGAAATTAATAGTGGCTATAAAGCTTATCGAGCCAAATACCTACTGTCCAAAG 1500  
DB 1488 CGATGCAAGAAATTAATAGTGGCTATAAAGCTTATCGAGCCAAATACCTACTGTCCAAAG 1547  
QY 1501 TCAGATGTGGATATTTTGGCGAGAGGTGCCATTTCTCTGCCAGCTCAATCATCCCTGC 1560  
DB 1548 TCAGATGTGGATATTTTGGCGAGAGGTGCCATTTCTCTGCCAGCTCAATCATCCCTGC 1607  
QY 1561 GTAAATTCAGTTTGGGTCCTTGTGTTGAATGATCCAGCCAGTTTGGCAATGTGCACTCAA 1620  
DB 1608 GTAAATTCAGTTTGGGTCCTTGTGTTGAATGATCCAGCCAGTTTGGCAATGTGCACTCAA 1667

QY 1621 TACATATCAGGGGTTCTCTGTTCTCCCTTCATGAGCAGAGAGATTCCTGATTG 1680  
Db 1668 TACATATCAGGGGTTCTCTGTTCTCCCTTCATGAGCAGAGAGATTCCTGATTG 1727  
QY 1681 CAGTCTAAATTAATTTAGAGTAGATGTTGCCAAGGATGAGTACCTTCACACCTG 1740  
Db 1728 CAGTCTAAATTAATTTAGAGTAGATGTTGCCAAGGATGAGTACCTTCACACCTG 1787  
QY 1741 ACACAGCAATATATACATCGTACTTGAAACAGTCAATATTTCTTCTATGAGATGG 1800  
Db 1788 ACACAGCAATATATACATCGTACTTGAAACAGTCAATATTTCTTCTATGAGATGG 1847  
QY 1801 CATGCTGGTGGGAGATTTTGGAGATCAAGATTTCTCAGTCTCTGGATGAAGACAC 1860  
Db 1848 CATGCTGGTGGGAGATTTTGGAGATCAAGATTTCTCAGTCTCTGGATGAAGACAC 1907  
QY 1861 ATGACAAACAACTGGGAACCTCCGTTGGATGGCTCTGAGGTGTTCAACGAGTGCAT 1920  
Db 1908 ATGACAAACAACTGGGAACCTCCGTTGGATGGCTCTGAGGTGTTCAACGAGTGCAT 1967  
QY 1921 CGGTACACCATCAAGCAGATGCTTCAGCTATGCTCTGTCTGTGGAAATTCCTACT 1980  
Db 1968 CGGTACACCATCAAGCAGATGCTTCAGCTATGCTCTGTCTGTGGAAATTCCTACT 2027  
QY 1981 GCGCAATTCATTCGCTCATCTCAAGCCAGCGCTCGGCGAGCAGACATGCTTACCAC 2040  
Db 2028 GCGCAATTCATTCGCTCATCTCAAGCCAGCGCTCGGCGAGCAGACATGCTTACCAC 2087  
QY 2041 CACATCAGACTCCCATGGCTATTCATTCGCAAGCCCATATCATCTCTGCTGATAGA 2100  
Db 2088 CACATCAGACTCCCATGGCTATTCATTCGCAAGCCCATATCATCTCTGCTGATAGA 2147  
QY 2101 GGTGGACGATGCTCTGAGGAGACCGAATTTCTGAAGTGTCTCATGAAGTAA 2160  
Db 2148 GGTGGACGATGCTCTGAGGAGACCGAATTTCTGAAGTGTCTCATGAAGTAA 2207  
QY 2161 GAGTGTCTCTCAACATGAGTGTCTCTGATCAAGTAAACAGAGTGGTCTCTC 2220  
Db 2208 GAGTGTCTCTCAACATGAGTGTCTCTGATCAAGTAAACAGAGTGGTCTCTC 2267  
QY 2221 TCACCTCTCTCTCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2280  
Db 2268 TCACCTCTCTCTCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2327  
QY 2281 GCAGATTAAGAAGTCTGTTTGAATGGAATATGCTCTAAATCAAGGTCCTATGCTGCT 2340  
Db 2328 GCAGATTAAGAAGTCTGTTTGAATGGAATATGCTCTAAATCAAGGTCCTATGCTGCT 2387  
QY 2341 TTGTCCCAAGTCTGACATATTCCTCTCAAGGTCTGCTTTGGAGGATGAAGA 2400  
Db 2388 TTGTCCCAAGTCTGACATATTCCTCTCAAGGTCTGCTTTGGAGGATGAAGA 2447  
QY 2401 AGTCTTCAATACACACCATGACAAATATGCTATGCTATCCGATCCCATGAGCTCAATG 2460  
Db 2448 AGTCTTCAATACACACCATGACAAATATGCTATGCTATCCGATCCCATGAGCTCAATG 2507  
QY 2461 CATTTTCATTTCCGAAATAGTAGAGTCTTTGAGACAGCAGC 2505  
Db 2508 CATTTTCATTTCCGAAATAGTAGAGTCTTTGAGACAGCAGC 2552

RESULT 4  
ACC48580  
ID ACC48580 standard; cDNA; 3025 BP.  
XX  
AC ACC48580;  
XX  
DT 11-AUG-2003 (first entry)  
XX  
DE Human cardiac-related ankyrin-repeat protein kinase (CARP) cDNA.  
XX  
KW Human; CARP; cardiac-related ankyrin-repeat protein kinase; enzyme;  
KW cardiant; hypotensive; cytostatic; chromosome 1; gene therapy; gene; ss.

XX OS Homo sapiens.  
XX FH Key Location/Qualifiers  
XX CDS 48..2555  
FT /\*tag= a  
FT /product= "Human CARP"  
FT /note= "the CDS is also claimed in Claim 1"  
XX  
XX MO2003020912-A2.  
XX  
XX PD 13-MAR-2003.  
XX  
XX 04-SEP-2002; 2002WO-US028300.  
XX  
XX 05-SEP-2001; 2001US-00947199.  
XX  
XX (MILL-) MILLENNIUM PHARM INC.  
XX  
XX Raju J;  
XX  
XX WPI; 2003-230188/28.  
XX  
XX Novel isolated cardiac-related ankyrin-repeat protein kinase polypeptide,  
XX useful for treating cellular growth related disorders which include  
XX cardiovascular disorders and proliferative and/or differentiative  
XX disorders.  
XX  
XX Claim 1; Fig 1a; 158pp; English.

The present sequence, the coding region of which is also claimed, is that of cDNA encoding a novel human cardiac-related ankyrin-repeat protein kinase, designated CARP, which plays a role in signalling pathways associated with cardiac cell growth or differentiation. The cDNA was isolated from a library prepared from subjects suffering from congestive heart failure of ischaemic and idiopathic origin. A plasmid containing the human CARP nucleotide sequence is deposited as ATCC PTA-1530. The CARP gene maps to chromosome 1, within the atrioventricular canal defects (AVCD) locus. The gene is therefore a candidate for a congenital heart defect susceptibility gene. CARP gene expression is increased in ischaemic heart tissue samples, further suggesting a role in the regulation of cardiac cell growth and/or differentiation and the pathogenesis of cardiovascular disorders, e.g. congestive heart failure and cardiac hypertrophy. The invention provides CARP nucleic acids, antisense molecules, recombinant expression vectors, host cells and transgenic animals in which a CARP gene has been introduced or disrupted, and also CARP proteins, fusion proteins, antigenic peptides and anti-CARP antibodies. These are useful as modulating agents for regulating a variety of cellular processes, e.g. cardiac cellular process, for modulating the phosphorylation state of a CARP molecule or one or more proteins involved in cellular growth or differentiation, for modulating cell behaviour or as targets and therapeutic agents controlling cardiac cell proliferation, differentiation, hypertrophy and migration, for modulating intra- or inter-cellular signalling and/or gene transcription, for modulating cell proliferation, growth, differentiation, survival and/or migration, for regulating transmission of signals from cellular receptors, for modulating entry of cells, e.g. cardiac precursor cells, into mitosis, or for regulating cytoskeletal function. The nucleic acids and proteins are useful for treating cellular growth related disorders which include cardiovascular disorders (such as heart failure, hypertension), and proliferative and/or differentiative disorders (such as cancer). They are also useful in screening assays, detection assays (e.g. chromosomal mapping, tissue typing, forensic biology), predictive medicine (e.g., diagnostic assays, prognostic assays, monitoring clinical trials) and pharmacogenomics, and in methods of treatment. The nucleic acids are also useful as probes and primers, in gene therapy. The nucleic detection of genetic alteration in the CARP gene, and in forensic identification

XX Sequence 3025 BP; 882 A; 622 C; 655 G; 866 T; 0 U; 0 Other;  
XX  
XX Query Match 100.0%; Score 2505; DB 7; Length 3025;  
XX Best Local Similarity 100.0%; Pred. No. 0;

Matches 2505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
Qy	1	ATGGGAAATTAATAATCTAGACCAACCCAAACTTGTACTGATGAATGGAAGAAAGATC	60	
Db	48	ATGGGAAATTAATAATCTAGACCAACCCAAACTTGTACTGATGAATGGAAGAAAGATC	107	
Qy	61	AGTGAATCATATGTTATCAATAGAAAGATTAGAGATGACCTCAGATCAAGGAAAAA	120	
Db	108	AGTGAATCATATGTTATCAATAGAAAGATTAGAGATGACCTCAGATCAAGGAAAAA	167	
Qy	121	GAATGACAGAACTAAGGAATATTTGGCTCTGATGAAGCCCTTCACTAAAGTCAATTTA	180	
Db	168	GAATGACAGAACTAAGGAATATTTGGCTCTGATGAAGCCCTTCACTAAAGTCAATTTA	227	
Qy	181	AAATTACGCACTGAAATGGCTCTCTACCTTCAATTTATGTTGATTTGGGGGCAAG	240	
Db	228	AAATTACGCACTGAAATGGCTCTCTACCTTCAATTTATGTTGATTTGGGGGCAAG	287	
Qy	241	AAATCACATATTCGAACCTTTATGTTGAAGGGCTCCGCCATCTCGACTGACAGAAAT	300	
Db	288	AAATCACATATTCGAACCTTTATGTTGAAGGGCTCCGCCATCTCGACTGACAGAAAT	347	
Qy	301	GGATTTACAGCTTGCATTTAGCAGTTTACAGGATAATGCGAATTTGATCACTTCTCTG	360	
Db	348	GGATTTACAGCTTGCATTTAGCAGTTTACAGGATAATGCGAATTTGATCACTTCTCTG	407	
Qy	361	CTTCACAGTGGAGCTGATATACAGCAGTTGGATACGGTGGCTCAGTCCCTCCATATT	420	
Db	408	CTTCACAGTGGAGCTGATATACAGCAGTTGGATACGGTGGCTCAGTCCCTCCATATT	467	
Qy	421	GCTACAAATAGCTGGCACTAGAGCTGCTGATGTGCTGTTGCAACATGGAGCTTAATGTC	480	
Db	468	GCTACAAATAGCTGGCACTAGAGCTGCTGATGTGCTGTTGCAACATGGAGCTTAATGTC	527	
Qy	481	AAATTTCAAGATCAGTTTCTTCTTCACTCCATTCGATTTGCGAGCTACTATGGACATGAA	540	
Db	528	AAATTTCAAGATCAGTTTCTTCTTCACTCCATTCGATTTGCGAGCTACTATGGACATGAA	587	
Qy	541	CAGTAACTCGCCTCTCTTTGAAATTTGGTGTGATGTAATGTAAGTGGTGAAGTTGGA	600	
Db	588	CAGTAACTCGCCTCTCTTTGAAATTTGGTGTGATGTAATGTAAGTGGTGAAGTTGGA	647	
Qy	601	GATAGACCCCTCACCTAGCATCTGCAAGGATTTCTTGAATTTGCAAACTCTTGATG	660	
Db	648	GATAGACCCCTCACCTAGCATCTGCAAGGATTTCTTGAATTTGCAAACTCTTGATG	707	
Qy	661	GAAGAGGCGACGAAGCAGATGTAATGCTCAAGATAATGAAGACATGTCCCACTCCAT	720	
Db	708	GAAGAGGCGACGAAGCAGATGTAATGCTCAAGATAATGAAGACATGTCCCACTCCAT	767	
Qy	721	TTCTGTTCTGATTTGGACACCATGATAGTTAAGTATCTGCTGCAAAAGTGAATTGGA	780	
Db	768	TTCTGTTCTGATTTGGACACCATGATAGTTAAGTATCTGCTGCAAAAGTGAATTGGA	827	
Qy	781	GTTCAACCTCATGTTGTAATTTCTATGAGATACCCCTTACCTGCGATGCTACAAAT	840	
Db	828	GTTCAACCTCATGTTGTAATTTCTATGAGATACCCCTTACCTGCGATGCTACAAAT	887	
Qy	841	GGCAAAATTTGAAGTTGCGAAGGAAATCATCCAAATATCAGGAACGAAAGTCTGACTAAG	900	
Db	888	GGCAAAATTTGAAGTTGCGAAGGAAATCATCCAAATATCAGGAACGAAAGTCTGACTAAG	947	
Qy	901	GAAACATCTTCAGTGAACAGCTTTTCTATGAGATACCCCTTACCTGCGATGCTACAAAT	960	
Db	948	GAAACATCTTCAGTGAACAGCTTTTCTATGAGATACCCCTTACCTGCGATGCTACAAAT	1007	
Qy	961	CTAGTCAAAATTTCTTCTTGTATCAGATGTCAATAACATCAACACCAAGGAGGATGGG	1020	
Db	1008	CTAGTCAAAATTTCTTCTTGTATCAGATGTCAATAACATCAACACCAAGGAGGATGGG	1067	
Qy	1021	CACACTGGATTAACCTGCTGCTACACGGTCAATTCGGCTGGTTCAGTTCTTACTG	1080	
Db	1068	CACACTGGATTAACCTGCTGCTACACGGTCAATTCGGCTGGTTCAGTTCTTACTG	1127	

Qy	1081	GATAATGGAGCTGATATGAATCTAGTGGCTTGTGATCCAGCAGGTCTAGTGGTGAATAA	1140	
Db	1128	GATAATGGAGCTGATATGAATCTAGTGGCTTGTGATCCAGCAGGTCTAGTGGTGAATAA	1187	
Qy	1141	GATGACACACATGTTTGTATGTTGGCTTATGAAAAGGGCATGATGCCATTTGTCACTC	1200	
Db	1188	GATGACACACATGTTTGTATGTTGGCTTATGAAAAGGGCATGATGCCATTTGTCACTC	1247	
Qy	1201	CTGAAGCATTTAAGAGACCAAGATGAATGGCTCTGAATGAATATCTCAGCCTGGA	1260	
Db	1248	CTGAAGCATTTAAGAGACCAAGATGAATGGCTCTGAATGAATATCTCAGCCTGGA	1307	
Qy	1261	GGAGATGGCTCTATGCTCTCCATCACCCCTTGGGGAAGATTAAAGCATGACAAAA	1320	
Db	1308	GGAGATGGCTCTATGCTCTCCATCACCCCTTGGGGAAGATTAAAGCATGACAAAA	1367	
Qy	1321	GAGAAGCAGATATTTCTCTCTTAAGAGCTGGATGGCTTCAATTTCCATCTTCAGCTC	1380	
Db	1368	GAGAAGCAGATATTTCTCTCTTAAGAGCTGGATGGCTTCAATTTCCATCTTCAGCTC	1427	
Qy	1381	TCAGAAATTTGAGTTCCATGAGATTATGGCTCAGTCTCTTTTGGGAAAGTATATAAGGA	1440	
Db	1428	TCAGAAATTTGAGTTCCATGAGATTATGGCTCAGTCTCTTTTGGGAAAGTATATAAGGA	1487	
Qy	1441	CGATGCAAGAAATAAATAGTGGCTATAAAACGTTATCGAGCCAAATACCTAGTCTCAAAG	1500	
Db	1488	CGATGCAAGAAATAAATAGTGGCTATAAAACGTTATCGAGCCAAATACCTAGTCTCAAAG	1547	
Qy	1501	TCAGATGTCGATATGTTTGGCGAGAGTGCCATTTCTGCCAGCTCAATCATCCCTGTC	1560	
Db	1548	TCAGATGTCGATATGTTTGGCGAGAGTGCCATTTCTGCCAGCTCAATCATCCCTGTC	1607	
Qy	1561	GTAATTCAGTTTGTGGTGTCTTGTGATGATCCAGCCAGTTTGGCAATTTGTCACTCAA	1620	
Db	1608	GTAATTCAGTTTGTGGTGTCTTGTGATGATCCAGCCAGTTTGGCAATTTGTCACTCAA	1667	
Qy	1621	TACATATCAGGGGGTCTCTGTTCTCCCTCTCTATGAGCAGAGAGATTTCTTGATTTG	1680	
Db	1668	TACATATCAGGGGGTCTCTGTTCTCCCTCTCTATGAGCAGAGAGATTTCTTGATTTG	1727	
Qy	1681	CAGTCTAAATTAATTTATTCAGTAGATTTGGCAAGCATGGAGTACTTCAACAACCTG	1740	
Db	1728	CAGTCTAAATTAATTTATTCAGTAGATTTGGCAAGCATGGAGTACTTCAACAACCTG	1787	
Qy	1741	ACACAGCCAAATTAACATCGTCACTTGAACAGTCAATATTTCTCTCTATGAGGATGGG	1800	
Db	1788	ACACAGCCAAATTAACATCGTCACTTGAACAGTCAATATTTCTCTCTATGAGGATGGG	1847	
Qy	1801	CATGCTGTGTCGCGAGATTTTGGAGATCAAGATTTCTACAGTCTCTGATGAAGACAAC	1860	
Db	1848	CATGCTGTGTCGCGAGATTTTGGAGATCAAGATTTCTACAGTCTCTGATGAAGACAAC	1907	
Qy	1861	ATGACAAAAACACCTGGGAACCTCCGTTGGATGGCTCTCTGAGGTGTTTCCAGCAGTGCACT	1920	
Db	1908	ATGACAAAAACACCTGGGAACCTCCGTTGGATGGCTCTCTGAGGTGTTTCCAGCAGTGCACT	1967	
Qy	1921	CGGTACACCATCAAGCAGATGTTTCACTGATGCTCTGTTCTGTTGGGAAATTTCTCACT	1980	
Db	1968	CGGTACACCATCAAGCAGATGTTTCACTGATGCTCTGTTCTGTTGGGAAATTTCTCACT	2027	
Qy	1981	GGGGAATTTCCATTCGCTCATCTCAAGCCAGGGGTGGGCGAGCAGACATGGCTTACCAC	2040	
Db	2028	GGGGAATTTCCATTCGCTCATCTCAAGCCAGGGGTGGGCGAGCAGACATGGCTTACCAC	2087	
Qy	2041	CACATCAGACCTCCCATTTGGCTTATTTCCATTTCCAAAGCCATATCATCTCTGCTGATAGA	2100	
Db	2088	CACATCAGACCTCCCATTTGGCTTATTTCCATTTCCAAAGCCATATCATCTCTGCTGATAGA	2147	
Qy	2101	GGGTGAAACGCTGCTCAAGAGAGACCCGAAATTTTCTGAGTTGTCTCATGAAGTTAGAA	2160	
Db	2148	GGGTGAAACGCTGCTCAAGAGAGACCCGAAATTTTCTGAGTTGTCTCATGAAGTTAGAA	2207	

QY 2161 GAGTGTCTCTCAACATTGAGCTGATGTCTCTGTCATCAAGTAAACAGCAGTGGTCTCTC 2220  
DB |||||  
QY 2208 GAGTGTCTCTCAACATTGAGCTGATGTCTCTGTCATCAAGTAAACAGCAGTGGTCTCTC 2267  
DB |||||  
QY 2221 TCACCTTCTCTCTCTCTGATTGCTGCTGAGTGAACCGGGAGGACCTGCGGAGTCATG 2280  
DB |||||  
QY 2268 TCACCTTCTCTCTCTCTGATTGCTGCTGAGTGAACCGGGAGGACCTGCGGAGTCATG 2327  
DB |||||  
QY 2281 GCAGCATTAAGAAGTCGTTTCCGAATTGGAAATATGCTCTAAATGCAAGGTCCTATGCTGCT 2340  
DB |||||  
QY 2328 GCAGCATTAAGAAGTCGTTTCCGAATTGGAAATATGCTCTAAATGCAAGGTCCTATGCTGCT 2387  
DB |||||  
QY 2341 TTGTCCTCAATGCTGAGCAATATTCCTCTCAAGGTCCTGCTTTGGAGGAGATGAAAAGA 2400  
DB |||||  
QY 2388 TTGTCCTCAATGCTGAGCAATATTCCTCTCAAGGTCCTGCTTTGGAGGAGATGAAAAGA 2447  
DB |||||  
QY 2401 AGTCTTCAATGCTGAGCAATATTCCTCTCAAGGTCCTGCTTTGGAGGAGATGAAAAGA 2460  
DB |||||  
QY 2448 AGTCTTCAATGCTGAGCAATATTCCTCTCAAGGTCCTGCTTTGGAGGAGATGAAAAGA 2507  
DB |||||  
QY 2461 CATTTTCATTTCCGAAATAGTAGCAGCTTTGAGGACAGCAGC 2505  
DB |||||  
QY 2508 CATTTTCATTTCCGAAATAGTAGCAGCTTTGAGGACAGCAGC 2552  
DB |||||

## RESULT 5

AAS80720  
ID AAS80720 standard; cDNA; 2788 BP.

AC AAS80720;

DT 13-FEB-2002 (first entry)

DE DNA encoding novel human diagnostic protein #16524.

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

OS Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US008631.

XX 31-MAR-2000; 2000US-00540217.

XX 23-AUG-2000; 2000US-00649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

XX P-PSDB; ASG16533.

XX New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity.

PS Claim 1; SEQ ID NO 16524; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)  
CC sequences. (I) is useful as hybridisation probes, polymerase chain  
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping.  
CC and in recombinant production of (II). The polynucleotides are also used  
CC in diagnostics as expressed sequence tags for identifying expressed  
CC genes. (II) is useful in gene therapy techniques to restore normal  
CC activity of (II) or to treat disease states involving (II). (II) is  
CC useful for generating antibodies against it, detecting or quantitating a  
CC polypeptide in tissue, as molecular weight markers and as a food  
CC supplement. (II) and its binding partners are useful in medical imaging

CC of sites expressing (II). (I) and (II) are useful for treating disorders  
CC involving aberrant protein expression or biological activity. The  
CC polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic  
CC coding sequences of the invention. Note: The sequence data for this  
CC patent did not appear in the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences

XX  
SQ Sequence 2788 BP; 807 A; 585 C; 626 G; 770 T; 0 U; 0 Other;

Query Match 98.0%; Score 2455.4; DB 5; Length 2788;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2467; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 38 CTGATGATCGAAGAAAGTCAAGTCAATGATGTTATCATCAATAGAAAGATTAGAG 97  
DB |||||  
QY 317 CTGATGATCGAAGAAAGTCAAGTCAATGATGTTATCATCAATAGAAAGATTAGAG 376  
DB |||||  
QY 98 ATGACCTGCGAGATCAAGGAAAGAACTGACAGAACTAAGGAATATATTGGCTCTGATG 157  
DB |||||  
QY 377 ATGACCTGCGAGATCAAGGAAAGAACTGACAGAACTAAGGAATATATTGGCTCTGATG 436  
DB |||||  
QY 158 AAGCTTTCAGTAAAGTCAATTTAAATACCGCACTGAAAAGGGCTGCTCTACTTCATT 217  
DB |||||  
QY 437 AAGCTTTCAGTAAAGTCAATTTAAATACCGCACTGAAAAGGGCTGCTCTACTTCATT 496  
DB |||||  
QY 218 TATGTTGCATTTGTGGAGGCAAGAAATCACATATTCGAACCTCTTATGTTGAAA-GGGCTC 276  
DB |||||  
QY 497 TATGTTGCATTTGTGGAGGCAAGAAATCACATATTCGAACCTCTTATGTTGAAAAGGGCTC 556  
DB |||||  
QY 277 CGCCCATCTCGACTGCAAGAAATGGAATTAACAGCTTGCATTTACAGAGTTTACAAGGAT 336  
DB |||||  
QY 557 CGCCCATCTCGACTGCAAGAAATGGAATTAACAGCTTGCATTTACAGAGTTTACAAGGAT 616  
DB |||||  
QY 337 AATGCGAAGTTGATCATTCTCTGCTTACAGTGGAGCTGATATACACAGCTTGATAC 396  
DB |||||  
QY 617 AATGCGAAGTTGATCATTCTCTGCTTACAGTGGAGCTGATATACAGAGCTTGATAC 676  
DB |||||  
QY 397 GGTGGCTCACTGCCCTCCCATATTGCTACAATAGCTGGCCACCTAGAGGCTGCTGATG 456  
DB |||||  
QY 677 GGTGGCTCACTGCCCTCCCATATTGCTACAATAGCTGGCCACCTAGAGGCTGCTGATG 736  
DB |||||  
QY 457 CTGTTGCAACATGGAGCTAATGTCATATTTCAAGTGCAGTTTCTTCTCACTCCATGAT 516  
DB |||||  
QY 737 CTGTTGCAACATGGAGCTAATGTCATATTTCAAGTGCAGTTTCTTCTCACTCCATGAT 796  
DB |||||  
QY 517 ATTGCGAGCTTACTATGGACATGAACAGGTAACTCCGCTTCTTTTGAATTTGGTCTGAT 576  
DB |||||  
QY 797 ATTGCGAGCTTACTATGGACATGAACAGGTAACTCCGCTTCTTTTGAATTTGGTCTGAT 856  
DB |||||  
QY 577 GTAATGTAAGTGGTGAAGTTGAGATAGACCCCTCCACCTAGCATCTGCAAAAGGATTC 636  
DB |||||  
QY 857 GTAATGTAAGTGGTGAAGTTGAGATAGACCCCTCCACCTAGCATCTGCAAAAGGATTC 916  
DB |||||  
QY 637 TTGAATATTGCAAAACTCTTTGATGGAAGGAGGAGCAAGAGAGATGTGAATGCTCAAGAT 696  
DB |||||  
QY 917 TTGAATATTGCAAAACTCTTTGATGGAAGGAGGAGCAAGAGAGATGTGAATGCTCAAGAT 976  
DB |||||  
QY 697 AATGAAGACCATGTCCTCCACTCCCATTTCTGTTCTCGATTTGGACACCATGATATAGTTAAG 756  
DB |||||  
QY 977 AATGAAGACCATGTCCTCCACTCCCATTTCTGTTCTCGATTTGGACACCATGATATAGTTAAG 1036  
DB |||||  
QY 757 TATCTCTGCAAGTCAATTTGGAAGTTCAACCTCATGTTGTTTAAATATCTATGGAGATACC 816  
DB |||||  
QY 1037 TATCTCTGCAAGTCAATTTGGAAGTTCAACCTCATGTTGTTTAAATATCTATGGAGATACC 1096  
DB |||||  
QY 817 CCCTTACACCTGGCATGCTACAAATGGCAAAATTTGAAGTTGCCAAGGAAATCATCAATA 876  
DB |||||  
QY 1097 CCCTTACACCTGGCATGCTACAAATGGCAAAATTTGAAGTTGCCAAGGAAATCATCAATA 1156  
DB |||||



XX Raju J;  
PI WPI; 2000-431275/37.  
DR  
XX New polynucleotide encoding cardiac-related ankyrin-repeat protein  
PT kinase, useful for treating disorders such as cardiovascular disorders,  
PT e.g. heart failure and cell differentiation disorders, e.g. cancer.  
XX  
PS Claim 1; Page 155-159; 161pp; English.  
XX  
CC CARX polypeptides are regulators of signal transmission from cellular  
CC receptors (e.g. cardiac cell growth factor receptor) and cytoskeleton  
CC function. They also act as modulators of cellular differentiation and  
CC cell death. Nucleic acids, or their fragments encoding CARX polypeptides  
CC are useful for detecting CARX nucleic acids especially mRNA, in a sample.  
CC CARX is useful for treating disorders associated with upregulation or  
CC downregulation of cellular proliferation such as, cardiovascular  
CC disorders (heart failure and hypertension) and disorders associated with  
CC cell differentiation such as cancer and sarcoma  
XX  
SQ Sequence 2505 BP; 648 A; 640 C; 617 G; 600 T; 0 U; 0 Other;  
Query Match 75.5%; Score 1892.2; DB 3; Length 2505;  
Best Local Similarity 84.7%; Pred. No. 0;  
Matches 2122; Conservative 0; Mismatches 383; Indels 0; Gaps 0;  
QY 1 ATGGGAATTATAATCTAGACCAACCCAACTTGTTACTGATGAATGGGAAGAAAGTC 60  
DB 1 ATGGGAATTACAAATCCAGACCAACACAGACTTGTTCTGATGAATGGGAAGAAAGTT 60  
QY 61 AGTGAATCATATGTTATACAAATAGAAAGATTAGAGATGACCTGCAGATCAAGGAATA 120  
DB 61 AGTGAATCTTACGCTTATATCATAGAAAGGCTGGAGGATAACCTGCAGATCAAGAAAT 120  
QY 121 GAACCTGACAGAACTAAGGAATATATTTGGCTCTGATGAAGCCTTCAGTAAAGTCAATTTA 180  
DB 121 GAATTTCAAGAACTAAGGCACATCTTTGGCTCTGATGAAGCCTTCAGTAAAGTCAATTTA 180  
QY 181 AATTACCGCATGAATAGGCTGTCTACTTCAATTTATGTTGCAATTTGTGGAGCAAG 240  
DB 181 AATTACCGCACAGACGGTGGCTCTCCCTCTGACACCTGTGCTGTCTGTGGCGGCAAC 240  
QY 241 AAATCACATATTCGAATCTTATGTTGAAAGGGCTCGGCCATCTGCACTGACAGAAAT 300  
DB 241 AAGTCACATATCCGTGGCTTATGTTTAAAGGGCTCCGTCCATCCAGACTGACGAGAAAT 300  
QY 301 GGATTTACAGCCTTGCATTTAGCAGTTTACAAGGAATATGCAAGATTTGATCTCTGT 360  
DB 301 GGGTTTCCAGCTCTGCACCTGGCCGTTTACAAGGACAGCCCGGAACCTTATCACTTCACTG 360  
QY 361 CTTACAGTGGAGTGATATACAGAGTTGGATACGGTGGCTCACTGCCCTCCATATT 420  
DB 361 TTGCACAGCGAGCAGATGTTTCAGCAAGTGGGATACGGTGGCTTCACGCCCTCCACATA 420  
QY 421 GCTCAATAGCTGACACCCAGAGGCTGCAGAAAGTGTGCTACAACATGGGSCCAATGTG 480  
DB 421 AATATTGAAGTGCAGTTTTTTCATCCATTCGATATTCAGCGTACTATGACATGAA 540  
QY 481 AATGTTCAAGATGCGGCTTCTTCAACCCCACTGCACATTCGACCCCTACTATGGGACGAG 540  
QY 541 CAGGTAATCCGCCCTCTTTTGAATTTGGTGTGATGTAATTAAGTGGTGAAGTTGGA 600  
DB 541 CAGGTAACAGTGTCTTTTGAAGTGTGCTGATGTCATTAAGCGGTGAAGTTGGG 600  
QY 601 GATAGACCCCTCCACCTAGCATCTGCAAAAGGATCTTGAATTCGAAACTCTTGATG 660  
DB 601 GACAGGCTCTGCACCTGGGCTCTGCAAGGGGCTTCTTCAACATTTGTGAATCTCCCTGGTA 660  
QY 661 GAAGAAGGCACAAAGCAGATGTGAATGCTCAAGATAATGAAGACCATGTCCCACTCCAT 720  
DB

DB 661 GAAGAAGGGAGCAAGACAGATGTGAACGCTCAGGACAATGAAGACACACGCTCCTCTGCAC 720  
QY 721 TTCTGTTCTCGAATTTGGACACCATGATATAGTTAAGTATCTGCTGCAAAAGTGAATTTGAA 780  
DB 721 TTCTGTTCTCGAATTTGGACACCAATATAGTGAGCTACCTGCTCAGAGTGACTAGAG 780  
QY 781 GTTCAACCTCATGTTGTTAATATCTATGAGATACCCCTTACACTGCGCATGCTACAAAT 840  
DB 781 GTCCAGCCTCAGCTCATTAACATCTATGCTGACACTCTTTTGCACCTGGCATGCTACAAAT 840  
QY 841 GGCATAATTTCAAGTTGCCAAGGAATCATCCAAATATCAGGAACAGAAAGTCTGACTAAG 900  
DB 841 GGAATTTTGAAGTTGCCAAGGAATTTGCCAGGTAAACAGGAACCTGAAAGTCTGACTAAG 900  
QY 901 GAAACATCTTCAGTGAAGACAGCTTTTCAATAGTCTTGATGCTATGCGCAAGAGCAATTGAC 960  
DB 901 GAAACATCTTCAGCGAGACAGCTTTTCAATAGTCTTGATGCTATGCGCAAGAGCAATTGAC 960  
QY 961 CTAGTCAAAATTTCTTCTGATCAGATGTCATAAATCATCAACCAACCAAGGAGGATGGG 1020  
DB 961 CTGTCATAATTTCTTCTGATCAGATGCTGTGAACATTAACCAACCGAAGAGATGGG 1020  
QY 1021 CACACTGGATTACACTCTGCTTGTACCAAGCTGCATTCGCTGGTTCAGTTCTTACTG 1080  
DB 1021 CACAGGATTTGCACTCTGCTTGTACCAAGGCCATATCCGCTGCTGGTTCAGTTCTTACTT 1080  
QY 1081 GATAATGGAGCTGATATGAATCTAGTGGCTTGTGATCCCGAGAGCTTAGTGGTGAATAA 1140  
DB 1081 GATAATGGTGCAGATATGAATCTTGTGCTTGTGATCCCGAGAGCTTAGTGGTGAATAA 1140  
QY 1141 GATGAGCAGACATGTTTGTGCTTATGAAAGGGCATGATGCCATTTGTCCACTC 1200  
DB 1141 GATGAGCAGACATGTTTGTGCTTATGAGGCTTACGAGAAAGGACATGATGCCATTTGTACACTC 1200  
QY 1201 CTGAAGCATTTAAGAGACACCAAGATGAATTTGCCCTGTATGAATATTTCTCAGCCTGGA 1260  
DB 1201 CTGAAGCATTTAAGAGACACCCAGGAGAGCTGCCATGTAAGCAATTTATCCAGCCTGGA 1260  
QY 1261 GAGATGGCTCTTATGCTGTCTGATCCATCACCTTTGGGGAAGATTAAAGCATGACAAA 1320  
DB 1261 GAGATGGCTCTTATGCTGTCTGATCCCTTCCCTTCCCTTGGGCAAGATTAAAGCATGACAAA 1320  
QY 1321 GAGAAGCAGATATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1380  
DB 1321 GAGAAGCAGATATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1380  
QY 1381 TCAGAAATTTGAGTTCCATGAGATTTAGTGGCTCAGGTTCTTTTGGGAAAGTATATAAGGA 1440  
DB 1381 TCAGAAATTTGAGTTCCAGGATTTAGTGGCTCAGGTTCTTTTGGGAAAGTATATAAGGG 1440  
QY 1441 CGATGAGAAATATAATAGTGGCTATAAAGCTTATCAGGCAATACCTACTGCTCCAAAG 1500  
DB 1441 CGATGAGAAATATAATAGTGGCTATAAAGCTTATCAGGCAATACCTACTGCTCCAAAG 1500  
QY 1501 TCAGATGTGATATGTTTTCGAGAGGTTGCTCATCTCTGCGAGCTCAATCATCCCTGTC 1560  
DB 1501 TCAGAGCTGATATGTTTTCGCGAGAGGTTGCTCATCTCTGCGAGCTCAACACCCCTGTC 1560  
QY 1561 GTAAATTCAGTTTGTGGGCTGCTTGTGATGATCCAGCAGTTTGGCAATTTGCTCACTCAA 1620  
DB 1561 GTGGTTTCAGTTTGTGGGCTGCTTGTGATGATCCAGGCTCAGTTTGGCAATTTGCTCACTCAG 1620  
QY 1621 TACATATCAGGGGTTCTCTGTTCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1680  
DB 1621 TACATTTTCAGGAGGCTCTCTGTTCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1680  
QY 1681 CAGTCTAAATTAATTAATTCAGTAGATGTTGCCAAGGCTAGGATGCTCTTCAACAACCTG 1740  
DB 1681 CAGTCTAAATTAATTAATTCAGTAGATGTTGCCAAGGCTAGGATGCTCTTCAACAACCTG 1740  
QY 1741 ACACAGCCAAATTAATTAATTCAGTAGATGTTGCCAAGGCTAGGATGCTCTTCAACAACCTG 1800  
DB 1741 ACCAGCCAAATTAATTAATTCAGTAGATGTTGCCAAGGCTAGGATGCTCTTCAACAACCTG 1800

```
QY 1801 CATGCTGTGGTGGCAGATTTGGAGAAATCAAGATTTCTACAGTCTCTGGATGAAGACAAC 1860
Db 1801 CATGCTGTGGTGGCAGATTTGGAGAAATCAAGATTTCTACAGTCTCTGGATGAAGACAAC 1860
QY 1861 ATACAAAACACCTGGGAACCTCGTTGGATGGCTCTGAGTGTTCACGAGTGCAC 1920
Db 1861 ATACAAAACACCTGGGAACCTCGTTGGATGGCTCTGAGTGTTCACGAGTGCAC 1920
QY 1921 CGGTACACCATCAAAAGCAGATGCTTTCAGCTATGCTGTGTCTGTGGAAATTCCTCACT 1980
Db 1921 AGATACACCATCAAGGCTGATGCTTTCAGTACTCCCTGTGTGTGGAGCTCCTCACT 1980
QY 1981 GGGGAAATTCATTCCTCATCTCAAGCAGCGGCTGGGAGCAGAGATGCTGTACAC 2040
Db 1981 GGGGAAATTCATTCCTCATCTCAAGCAGCGGCTGGGAGCAGAGATGCTGTACAC 2040
QY 2041 CACATCAGACCTCCCATTCGGTATTCATTCCTCAAGCCCATATCTCTCTGATAGA 2100
Db 2041 CACATCAGACCTCCCATTCGGTATTCATTCCTCAAGCCCATCTCATCTCTGATACGG 2100
QY 2101 GGGTGAACGCGATGCTCTGAAGAGACCCGAAATTTCTGAAGTTGTCTAAGATTGAA 2160
Db 2101 GGGTGAATGCGATGCTCTGAAGAGACCCAGAGTTCTCTGAAGTCTGTAGCAAACTGGAG 2160
QY 2161 GAGTGTCTCTGCAACATTCAGTGTGATGCTCTCTGATCAAGTAAACAGAGTGGTCTCTC 2220
Db 2161 GAGTGTCTATGCAATGTGAGTCTATGCTCTCAGCATCAAGTAAACAGAGTGGTCTCTC 2220
QY 2221 TCACCTTCTTCTTCTCTGATTCGCTGTGATGCGGAGGAGACCTGGCGGAGTCACTGTG 2280
Db 2221 TCACCTTCTTCTTCTCTGATTCGCTGTGATGCGGAGGAGGCGCTGGCGGAGCACGTG 2280
QY 2281 GCAGCATTAAGAGTGTTCGAATTTGGAATGATGCTCTAAATGCAAGTCTCTATGCTGCT 2340
Db 2281 GCAGCCTTACGGAGCGGTTTGAATTTGAGTGGAGTATGCGCTAAATGCAAGTCTCTATGCTGGG 2340
QY 2341 TTGCTCCAAAGTCTGACAAATATTCCTCTCAAGTCTGTCTTTGGAGAGATGAAGA 2400
Db 2341 TGTCTCCAAAGTCTGGAACACACTCTATTCGGGCTGTCTTTGGAGAGATGAATAGG 2400
QY 2401 AGTCTTCAATACACCCCATTTGCAAAATATGCTATGATTCGATCCCATGAGTCTCAATG 2460
Db 2401 AGCACCCAGTATTCAACTGTTGACAAATACGGCTATGCTCTGATCCCATGAGCTGACG 2460
QY 2461 CATTTCATCTTTCGGAATAGTAGCAGCTTTGAGGACAGCAGC 2505
Db 2461 CACTTCACTTCCGCCCAAGACGACGAACTTTGAGGACAGCAAC 2505
```

## RESULT 7

AAA47608  
ID AAA47608 standard; cDNA; 3026 BP.

XX AC AAA47608;

XX 20-OCT-2000 (first entry)

DE Rat CARK (Cardiac related Ankyrin-Repeat Protein Kinase) cDNA.

XX Cardiac related ankyrin repeat protein kinase; CARK; cytoskeleton;  
KW cardiac cell growth factor receptor; cell differentiation; modulator;  
KW regulator; detection; cellular proliferation; cardiovascular disorder;  
KW heart failure; hypertension; cancer; sarcoma; ds.

OS Rattus norvegicus.

XX PH Key Location/Qualifiers  
FT CDS 61..2568

FT /\*tag= a  
FT /product= "Rat CARK"

XX WO200034330-A1.

```
XX 15-JUN-2000.  
XX 10-DEC-1999; 99WO-US029465.  
XX 11-DEC-1998; 98US-0111938P.  
XX 14-APR-1999; 99US-00291899.  
XX (MILL-) MILLENNIUM PHARM INC.  
XX Raju J;  
XX WPI; 2000-431275/37.  
XX P-PSDB; AAB01474.  
XX New polynucleotide encoding cardiac-related ankyrin-repeat protein  
PT kinase, useful for treating disorders such as cardiovascular disorders,  
PT e.g. heart failure and cell differentiation disorders, e.g. cancer.  
XX  
XX Claim 1; Fig 5; 151pp; English.  
XX CARK polypeptides are regulators of signal transmission from cellular  
CC receptors (e.g. cardiac cell growth factor receptor) and cytoskeleton  
CC function. They also act as modulators of cellular differentiation and  
CC cell death. Nucleic acids, or their fragments encoding CARK polypeptides  
CC are useful for detecting CARK nucleic acids especially mRNA, in a sample.  
CC CARK is useful for treating disorders associated with upregulation or  
CC downregulation of cellular proliferation such as, cardiovascular  
CC disorders (heart failure and hypertension) and disorders associated with  
CC cell differentiation such as cancer and sarcoma  
XX  
XX Sequence 3026 BP; 778 A; 751 C; 743 G; 754 T; 0 U; 0 Other;
```

Query Match 75.5%; Score 1892.2; DB 3; Length 3026;

Best Local Similarity 84.7%; Pred. No. 0;

Matches 2122; Conservative 0; Mismatches 383; Indels 0; Gaps 0;

```
QY 1 ATGGGAAATTTATAAATCTAGACCAACCCAACTTCTACTGATGATGGAAGAAAAGTC 60
Db 61 ATGGGAAATTTACAAATCCAGACCAACACAGACTTCTCTGATGATGGAAGAAAAGTT 120
QY 61 AGTGAATCATATGTTTATCAATGAAGAAATGAGATGACCTGCAGATCAAGGAAAAA 120
Db 121 AGTGAATCTTTAGCGTATTTATCATAGAAAGGCTGGAGGATAACCTGCAGATCAAGAAAAT 180
QY 121 GAATGACAGAACTAAGGAATATATTTGGCTCTGATGAAGCCTTCAGTAAAGTCAATTTA 180
Db 181 GAATTTCAAGAACTAAGGCAATCTTTGGCTCTGATGAAGCCTTCAGTAAATCAGTTTA 240
QY 181 AATTACCGCACTGAAAATGGGCTGTCTTACTTATTTATGTTGATTTGAGGAGCAAG 240
Db 241 AATTACCGCACAGAGCGTGGCGCTGTCCCTGTCTACACCTCTGTGTCTGTGGCGCAAC 300
QY 241 AATACATATTTGAACTCTTATGTTGAAGGGCTCCGCCATCTCGACTGACAGAAAT 300
Db 301 AAGTCACATATCCGTGCCCTTTAATGTTAAAGGGCTCCGTCCATCCAGACTGACAGAAAT 360
QY 301 GGATTTACAGCCTTGCAATTTAGCAGTTTACAAGGATAATGAGAAATGATCATTCTCTG 360
Db 361 GGGTTTCCAGCTCTGCACCTGSCCGTTTACAAGGACAGCCCGGAACTTATCATTCACTG 420
QY 361 CTTCAAGTGGAGCTGATATACAGAGGTTGATACGTTGAGTACGTTGCCTCACTGCCCTCATATT 420
Db 421 TTGCACAGCGGAGCAGATGTTTCAGCAAGTGGGATACCGTGGCCTCACAGCCCTCCACATA 480
QY 421 GCTACAATAGCTGGCCACCTAGAGGCTGTGATGTGCTGTTGCAACATGGAGTAAATGTC 480
Db 481 GCTGCAATAGCTGGACACCCAGAGGCTGCAAGATGCTGCTACAACATGGGGCCCAATGTG 540
QY 481 AATATTCAAGATGCGAGTGTCTTTTCTACTCCATTCGATATGAGCGGTACTATGGACATGAA 540
Db 541 AATGTTCAAGATGCGGTCTTCTTCTACCCCACTGCAATTTGACGCTTACTATGGGCACGAG 600
```

QY 541 CAGGTAACCTCCCTTCTTTGAAATTTGGTGTGCTGATGTAATGTAAGTGGTGAAGTTGGA 600  
DB 601 CAGGTAACCAAGTCCCTTTTGAAGTTTGGTGTGCTGATGTAATGTAAGCGGTGAAGTTGGG 660  
QY 601 GATAGACCCCTCCACCTTAGCATCTGCAAAAGGATTTCTTGAATATTGCAAAACTCTTTGATG 660  
DB 661 GACAGGCTCTGCACTGGCTCTGCAAGGGCTTTCTTCAACATTGTGAAACTCTCTGTA 720  
QY 661 GAAGAAGCAGCAAGCAGATGTAATGCTCAAGATTAATGAAGCAGATGCCATCTCCAT 720  
DB 721 GAAGAAGGAGCAAGCAGATGTAATGCTCAAGATTAATGAAGCAGATGCCATCTCCAT 780  
QY 721 TTCTGTTCTCCATTTGGACACCATGATATAGTTAAAGTATCTGCTGCAAAAGTATTGGA 780  
DB 781 TTCTGTTCTCCATTTGGACACCAATATATAGTGAAGTATCTGCTCCAGAGTACTTAGAG 840  
QY 781 GTTCAACCTCATGTTGTTAAATATCTATGGAGATACCCCTTTACACCTGGCATCTACAAT 840  
DB 841 GTCCAGCCTCACGTCATTAACTATATGTTGACACTCTTTTGACCTGGCATCTACAAT 900  
QY 841 GGCATAATTGAAGTTGCCAGGAATCATCCAATATACAGAAACAGAAAGTCTGACTAG 900  
DB 901 GGAATTTTGAAGTTGCCAGGAATTTGTCCAGGTAAACAGGAATCTGAAAGTCTGACTAG 960  
QY 901 GAAACATCTTCAGTGAACACAGCTTTTTCATAGTGTCTGTACCTATGGCAAGCATTTGAC 960  
DB 961 GAAACATCTTCAGGAGACAGCTTTTTCACAGTGTCTGTACCTATGGCAAGCATTTGAC 1020  
QY 961 CTAGTCAAAATTTCTTTGATCAGAAATGTCATAAATCAACCAAGCAAGAGGATGGG 1020  
DB 1021 CTGCTCAAAATTTCTTTGATCAGAAATGTCATAAATCAACCAAGCAAGAGATGGG 1080  
QY 1021 CACACTGATTAACACTCTGCTTGTCTACCAAGGTCAATTCGCTGCTGTTCACTTCTTACTG 1080  
DB 1081 CACACAGATTTGCACTCTGCTTGTCTACCAAGGCAATTCGCTGCTGTTCACTTCTTACTG 1140  
QY 1081 GATAATGGAGTGTATGAATCTAGTGGTGTGATCCCAAGAGTCTAGTGGTGAATAA 1140  
DB 1141 GATAATGGTGCAGATATGAATCTTGTGCTGTGATCCCAAGAGTCTAGTGGTGAATAA 1200  
QY 1141 GATGAGCAGACATGTTGATGTTGGCTTATGAAGAGGCAATGATCCATTTGCACTC 1200  
DB 1201 GATGAGCAGACATGTTGATGTTGGCTTATGAAGAGGCAATGATCCATTTGCACTC 1260  
QY 1201 CTGAAGCATTTAAGAGACCAACAGATGAATGCCCTGTAAATGAATATTTCTCAGCTGGA 1260  
DB 1261 CTGAAGCATTTAAGAGACCAACAGAGAGCTGCCATGTAAACGAATATTTCCAGCTGGA 1320  
QY 1261 GGAGATGGCTCTATGTTGTCTGTTCCATCACTTGGGAGAGATTAAGAGCATGACAAA 1320  
DB 1321 GGAGATGGCTCTATGTTGTCTGTTCCCTTGGGCAAGATTAAGAGCATGACAAA 1380  
QY 1321 GAGAAGGCAGATATTTCTCTCCCTAAGAGCTGGATTTGCCCTTCAATTTCCATCTCAGCTC 1380  
DB 1381 GAGAAGGCAGATATTTCTCTCCCTCAGGGCTGAACTACCTCCGCTTCCATCTCCACTC 1440  
QY 1381 TCAGAAATTTGATTTCCATGAGATTAATTTGGCTCAGGTTCTTTTGGGAAAGTATATAAGGA 1440  
DB 1441 TCCGAAATTCGAGTTCCAGAGATTAATCGGCTCGGGTTCTTTTGGGAAAGTCTATAAGGG 1500  
QY 1441 CGATGCGAATAAATAAGTGGCTATAAAGCTTTATCGAGCCCAATACCTACTGCTCCAG 1500  
DB 1501 CGATGCGAATAAATAAGTGGCTATAAAGCTTTATCGAGCCCAATACCTACTGCTCCAG 1560  
QY 1501 TCAGATGTGGATATTTTTCGAGAGAGTGTCCATTTCTGCTGCGAGTCAATCATCCCTGC 1560  
DB 1561 TCAGAGTGGATATGTTTTCGAGAGAGTGTCCATTTCTGCTGCGAGTCAATCATCCCTGC 1620  
QY 1561 GTAATTCAGTTTGGGTGCTGCTGATGATCCAGCCAGTTCGCCATTTGCACTCAA 1620  
DB 1621 GTGGTTTCAAGTTTGGGTGCTGCTGATGATCCAGTCAAGTTCGCCATTTGCACTCAA 1680  
QY 1621 TACATATCAGGGGGTCTCTGTTCTCTCCCTCTTTCATGAGCAGAAAGAGATTTCTGATTG 1680

DB 1681 TACATTTTCAGAGGCTCCCTGTTCTCCCTGCTTTCATGAACAGAAAGATTTCTTACTG 1740  
QY 1681 CAGTCTAAATTAATATTAGTGTAGATTTGGCAAGGATGAGATGAGTCTTCAACACTG 1740  
DB 1741 CAGTCTAAATTAATATTAGTGTAGATTTGGCAAGGATGAGATGAGTCTTCAACACTG 1800  
QY 1741 ACACAGCAATTAATATTAGTGTAGATTTGGCAAGGATGAGATGAGTCTTCAACACTG 1800  
DB 1801 ACCAGCAATTAATATTAGTGTAGATTTGGCAAGGATGAGATGAGTCTTCAACACTG 1860  
QY 1801 CATGCTGTGTGGCAGATTTTGGAGAAATCAAGATTTTCAAGTCTCTGATGAAGCAAC 1860  
DB 1861 CATGCTGTGTGGCAGATTTTGGAGAAATCAAGATTTTCAAGTCTCTGATGAAGCAAC 1920  
QY 1861 ATGACAAACCACTGGGAACCTCGTGGATGGCTCTGAGGTGTTACGAGTGCAT 1920  
DB 1921 ATGACAAAGCAGCAGGAACTCGCTGGATGGCTCTGAGGTGTTACGAGTGCAT 1980  
QY 1921 CGGTACACCATCAAGCAGATGTTTCAAGTATGCTCTGCTGTCTGTTGGAATTTCTCACT 1980  
DB 1981 AGATACACCATCAAGCAGTGTCTTCAAGTATGCTCTGCTGTCTGTTGGAATTTCTCACT 2040  
QY 1981 GGCAGAAATTCATTTGCTCATCAAGCAGGCTGTGGGAGAGATGAGTCTTACAC 2040  
DB 2041 GGCAGAAATTCATTTGCTCATCAAGCAGGCTGTGGGAGAGATGAGTCTTACAC 2100  
QY 2041 CACATCAGACCTCCCATTTGGCTATTTCCATTTCCCAAGCCCATATCTCTGCTGATACA 2100  
DB 2101 CACATCAGACCTCCCATTTGGCTATTTCCATTTCCCAAGCCCATATCTCTGCTGATACA 2160  
QY 2101 GGGTGAAGCAGATGTTCTTGAAGAGAGACCCGAAATTTCTGAAGTGTCTCAAGTTAGAA 2160  
DB 2161 GGGTGAAGTGCATGTTCTTGAAGAGAGACCCGAAATTTCTGAAGTGTCTCAAGTTAGAA 2220  
QY 2161 GAGTGTCTGCAACATTTGAGCTGTCTCTGATCAAGTCAAGTCAAGCAGTGGTCTCTC 2220  
DB 2221 GAGTGTCTGCAACATTTGAGCTGTCTCTGATCAAGTCAAGTCAAGCAGTGGTCTCTC 2280  
QY 2221 TCACCTTCTTCTTCTTCTGATTTGCTGTAACCCGAGGAGACCTGGCCGAGTCAATG 2280  
DB 2281 TCACCTTCTTCTTCTTCTGATTTGCTGTAACCCGAGGAGACCTGGCCGAGTCAATG 2340  
QY 2281 GCAAGCATTAAGAGTGTCTTGAATTTGGAATGATCTTAAATGCAAGTCTCTATGCTCT 2340  
DB 2341 GCAAGCATTAAGAGTGTCTTGAATTTGGAATGATCTTAAATGCAAGTCTCTATGCTCT 2400  
QY 2341 TTGTCCTCAAGTGTCTGCAATATTTCTCTCAAGTGTCTGTTTGGAGAGATGAAGA 2400  
DB 2401 TGTTCCTCAAGTGTCTGCAATATTTCTCTCAAGTGTCTGTTTGGAGAGATGAAGA 2460  
QY 2401 AGTCTTCAATACACCATTTGCAAAATGATGCTATGATTCGATTCGATCCCATGATG 2460  
DB 2461 AGCAGCAGTATTTCAACTGTTGCAAAATACGCTATGTTCTGATTCGATCCCATGATG 2520  
QY 2461 CATTTTCTTCTTCTGCGAATAGTAGCAGCTTTTGGAGCAGCAGC 2505  
DB 2521 CACCTTCACTCCCGCAAGACGACAGCAACTTTGAGGACAGCAAC 2565

## RESULT 8

ACC48581  
ID ACC48581 standard; cDNA; 3026 BP.

XX ACC48581;

DT 11-AUG-2003 (first entry)

XX Rat cardiac-related ankyrin-repeat protein kinase (CARK) cDNA.

XX Rat; CARK; cardiac-related ankyrin-repeat protein kinase; enzyme;  
KW cardiant; hypotensive; cytosolic; chromosome 1; gene therapy; gene; ss.



1141 GATGAGCAGCATGTTTGGCTTATGAAAAAGGCGATGATGCCATGTCACATC 1200  
Db  
1201 GATGAGCAGCATGTTTGGCTTATGAAAAAGGCGATGATGCCATGTCACATC 1260  
Qy  
1201 CTGAAGCAATTAAGAGACCAACAGATGAATGCCCTGTAATGAATATTCAGCCTGA 1260  
Db  
1261 CTGAAGCACTACAAGAGACCCAGAGAGCTGCATGTAAACGAATATTCAGCCTGA 1320  
Qy  
1261 GAGATGCTCCTATGTCGTCATCAACCTTGGGAGATTAAGAGCATGACAAA 1320  
Db  
1321 GAGATGCTCCTATGTCGTCATCAACCTTGGGAGATTAAGAGCATGACAAA 1380  
Qy  
1321 GAGAGGCGAGATATTCCTCCTAAGAGCTGGATTCCTTCAATTTCCATCTTCAGCTC 1380  
Db  
1381 GAGAGGCGAGATATTCCTCCTGAGGCTGAATACCTCCCGCTTCCATCTCCAATC 1440  
Qy  
1381 TCAGAAATGAGTTCATGAGATTAATGGCTCAGGTTCTTTGGGAAAGTATATAAGA 1440  
Db  
1441 TCCGAAATCGAGTTCAGAGATTAATCGGCTCGGGTTCCTTTGGGAAAGTCTATAAGG 1500  
Qy  
1441 CGATCAGAAATAAATAGTGGCTATAAAGCTTATCGAGCCAATACCTACTGCTCCAAG 1500  
Db  
1501 CGATCAGAAATAAATAGTGGCTATAAAGCTTATCGAGCCAATACCTACTGCTCCAAG 1560  
Qy  
1501 TCAGATGAGATATGTTTGGCGAGAGGTGCTATCTCTGCGAGCTCAATCATCCTCTC 1560  
Db  
1561 TCAGAGTGGATATGTTTGGCGAGAGGTGCTATCTCTGCGAGCTCAATCATCCTCTC 1620  
Qy  
1561 GTAAATTCAGTTTGGGCTGCTGCTGAATGATCCAGCCAGTTGCCATTTGCACTCAA 1620  
Db  
1621 GTGGTTGAGTTTGGGCTGCTGCTGAATGATCCAGCCAGTTGCCATTTGCACTCAA 1680  
Qy  
1621 TACATATCAGGGGTTCTGTTCTCCTCCTCTCATGAGCAGAGAGATTTCTGATTTG 1680  
Db  
1681 TACATTTGAGGAGTCTCCTGTTCTCCTCCTCATGAACAGAGAGATTTCTGATTTG 1740  
Qy  
1681 CAGTCTAAATTAATTTGAGTAGATGTTGCCAAGGATGAGTACCTTCAACCTG 1740  
Db  
1741 CAGTCTAAATTAATTTGAGTAGATGTTGCCAAGGATGAGTACCTTCAACCTG 1800  
Qy  
1741 ACACAGCAATTAATACATCGTGAATGAACAGTCAACATATTTCTCTATGAGGATGG 1800  
Db  
1801 ACCAGCAATCATACCGGACCTGAACAGCCACATATTTCTGCTATGAGGATGG 1860  
Qy  
1801 CATGCTGTTGGGAGATTTTGGAGATCAAGATTTCTACAGTCTCTGGATGAAGCAAC 1860  
Db  
1861 CATGCTGTTGGGAGATTTTGGAGATCAAGATTTCTGAGTCCCTGGATGAAGCAAC 1920  
Qy  
1861 ATGACAAACCACTGGGAACCTCCGTTGGATGGCTCCTGAGTGTTCACGAGTGCAT 1920  
Db  
1921 ATGACAAACCACTGGGAACCTCCGTTGGATGGCTCCTGAGTGTTCACGAGTGCAG 1980  
Qy  
1921 CGGTACACCATCAAGCAGATGCTTCAAGTATGCTCTGTTGTTGGGAAATTTCTACT 1980  
Db  
1981 AGATACACCATCAAGCAGATGCTTCAAGTATGCTCTGTTGTTGGGAAATTTCTACT 2040  
Qy  
1981 GGCGAAATTCATTCGCTCATCAAGCAGGCTGCGCAGCAGACATGCTTACAC 2040  
Db  
2041 GGAGAAATTCATTCGCTCATCAAGCAGGCTGCGCAGCAGACATGCTTACAC 2100  
Qy  
2041 CACATCAGACCTCCCATGCTATTCATTTCCCAAGCCCATATCATCTCTGCTGATAGA 2100  
Db  
2101 CACATCAGACCTCCCATGCTATTCATTTCCCAAGCCCATATCATCTCTGCTGATAGA 2160  
Qy  
2101 GGTGGAGCAGATGCTCTGAGGAGACCCGATTTCTGAGTTGCTGATGAGTAGNA 2160  
Db  
2161 GGTGGAGCAGATGCTCTGAGGAGACCCGATTTCTGAGTTGCTGATGAGTAGNA 2220  
Qy  
2161 CAGTGTCTCTCAACATGAGCTGATGCTCTCTGATCAAGTAAACAGAGTGGTCTCTC 2220  
Db  
2221 GAGTGGCTATGCAATGTTGGAGTCTATGCTTCCAGCATCAAGTAAACAGAGTGGTCTCTC 2280  
Qy  
2221 TCAGCTTCTCTTCTGATGCTGCTGGTGAACCGGGGAGGACCTGGCGGAGTATGATG 2280

2281 TCACCTTCTCTTCTTCCGATTTGCTGCTGAGCCGGAGGCGCTTGGCGGAGCACGTG 2340  
Qy  
2281 GCAGCATTAAAGAGTGTGTTTCCGATTTGGAATATGCTTAAATGCAAGTCTTATGCTCT 2340  
Db  
2341 GCAGCCTTACGGAGCCGTTTGGATTTGGAGTATGCCCTAAATGCAAGTCTTATGCTGG 2400  
Qy  
2341 TTGTCCCAAGTGTCTGCAAAATATCTCTCAAGTGTCTTCTTGGAGGAGATGAAAGA 2400  
Db  
2401 TGTCCCAAGTGTGGAACACACTCTAATCCGGGCTGTCTTTGAGGAGATGAATAGG 2460  
Qy  
2401 AGTCTTCAATACACCCATTTGCAAAATATGCTATGATCCGATCCCATGAGTCAATG 2460  
Db  
2461 AGCACCAGTATTAACACTGTTGCAAAATACGGCTATGTCTGTATGATCCCATGAGCCTGAG 2520  
Qy  
2461 CATTTTCATTTCTGCGGAAATAGTAGCAGCTTTGAGGACAGCAGC 2505  
Db  
2521 CACCTTCACTCCCGCAAGACGACGACCACTTTGAGGACAGCAAC 2565

RESULT 9  
AAD57334  
ID AAD57334 standard; cDNA; 2024 BP.  
XX  
AC AAD57334;  
XX  
DT 06-NOV-2003 (first entry)  
XX  
XX Human kinase and phosphatase (KPP-7) cDNA.  
DE Human; kinase; phosphatase; KPP; cell proliferative disorder; hepatitis;  
XX atherosclerosis; cirrhosis; haemoglobinuria; polycythaemia vera; cancer;  
KW psoriasis; thrombocytopenia; developmental disorder; Reiter's syndrome;  
KW renal tubular acidosis; anaemia; mental retardation; Alzheimer's disease;  
KW neurological disorder; Parkinson's disease; epilepsy; glomerulonephritis;  
KW autoimmune disorder; inflammatory disorder; AIDS; Goodpasture's syndrome;  
KW acquired immune deficiency syndrome; immunosuppressive; Crohn's disease;  
KW neutropenic; dermatitis; multiple sclerosis; diabetes mellitus;  
KW allergy; gout; Grave's disease; Hashimoto's thyroiditis; bowel syndrome;  
KW osteoporosis; rheumatoid arthritis; Sjogren's syndrome; ophthalmological;  
KW gene therapy; asthma; anticonvulsant; uropathic; pancreatitis; gene; ss.  
XX Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 46..1857  
FT /\*tag= a  
FT /product= "Human KPP"  
XX  
XX WO2003050084-A2.  
PD 19-JUN-2003.  
XX  
XX 06-DEC-2002; 2002WO-US039126.  
PR 07-DEC-2001; 2001US-0340235P.  
PR 19-DEC-2001; 2001US-0343007P.  
PR 21-DEC-2001; 2001US-034346P.  
PR 04-FEB-2002; 2002US-0354388P.  
PR 15-FEB-2002; 2002US-0357675P.  
XX  
XX (INCY-) INCYTE GENOMICS INC.  
XX  
XX Kable AE, Chien D, Wilson AD, Swarnakar A, Gorvad AE;  
PI Hafalia AJA, Emerling BM, Rankumar J, Jin P, Griffin JA, Marquis JP;  
PI Baughn MR, Chawla NK, Lehr-Mason PM, Khare R, Lee S, Hawkins PR;  
PI Becha SD, Lee SY, Sprague WW, Zebbarjadian Y;  
XX WPI; 2003-532894/50.  
DR P-PSDB; AAE37962.  
XX  
XX New human kinases and phosphatases and polynucleotides, useful for  
PT diagnosing, treating or preventing autoimmune or inflammatory disorders

PT (e.g. AIDS, allergy or anemia), multiple sclerosis, osteoarthritis,  
XX cancer or hepatitis.

XX Claim 5; Page 254; 282pp; English.

XX The invention relates to an isolated polypeptide, which is a human kinase  
CC and phosphatase (KPP). KPP agonists and antagonists are useful for  
CC diagnosing, treating or preventing disorders associated with aberrant  
CC expression of KPP, particularly cell proliferative disorders (e.g.  
CC arteriosclerosis, atherosclerosis, cirrhosis, hepatitis, paroxysmal  
CC nocturnal haemoglobinuria, polycythemia vera, psoriasis, primary  
CC thrombocytopaenia or cancer), developmental disorders (eg. renal tubular  
CC acidosis, anaemia or mental retardation), neurological disorders (e.g.  
CC Alzheimer's disease, Parkinson's disease or epilepsy), autoimmune/  
CC inflammatory disorders (e.g. AIDS; acquired immune deficiency syndrome,  
CC allergies, asthma, autoimmune thyroiditis, contact dermatitis, Crohn's  
CC disease, diabetes mellitus, glomerulonephritis, Goodpasture's syndrome,  
CC gout, Grave's disease, Hashimoto's thyroiditis, irritable bowel syndrome,  
CC multiple sclerosis, osteoarthritis, osteoporosis, pancreatitis, Reiter's  
CC syndrome, rheumatoid arthritis, Sjogren's syndrome, uveitis), or viral,  
CC bacterial, fungal, parasitic, protozoan or helminthic infections. The KPP  
CC is useful in assessing the effects of exogenous compounds on the  
CC expression of nucleic acids and kinases and phosphatases. KPP gene is  
CC useful in gene therapy and for creating transgenic animals to model human  
CC disease. The present sequence is human KPP cDNA

XX SQ Sequence 2024 BP; 613 A; 431 C; 425 G; 555 T; 0 U; 0 Other;

Query Match 70.7%; Score 1772; DB 8; Length 2024;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 1775; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 ATGGGAAATTATAAATTAGACCAACCCAACTTGATGATGGAAGGAAAGGTC 60  
DB 46 ATGGGAAATTATAAATTAGACCAACCCAACTTGATGATGGAAGGAAAGGTC 105  
QY 61 AGTGAATCATATGTTATCAATAGAAAGATTAGAAAGATGACTCAGATCAAGGAAAA 120  
DB 106 AGTGAATCATATGTTATCAATAGAAAGATTAGAAAGATGACTCAGATCAAGGAAAA 165  
QY 121 GAATGACAGAACTAAGGAATATTTGCTGTGATGAAGGCTTCAGTAAAGTCAATTTA 180  
DB 166 GAATGACAGAACTAAGGAATATTTGCTGTGATGAAGGCTTCAGTAAAGTCAATTTA 225  
QY 181 AATTACCGACTGAAATGGGCTGCTCTACTTCAATTTATGTTGCAATTTGAGGCAAG 240  
DB 226 AATTACCGACTGAAATGGGCTGCTCTACTTCAATTTATGTTGCAATTTGAGGCAAG 285  
QY 241 AATACATATTCGAATCTTATGTTGAAGGCTCCGCCATCTCGACTGACAGAAT 300  
DB 286 AATACATATTCGAATCTTATGTTGAAGGCTCCGCCATCTCGACTGACAGAAT 345  
QY 301 GGATTTACAGCCTTGCAATTTAGCAGTTTACAAGGATAATGCAAGATTGATCACTTCTCTG 360  
DB 346 GGATTTACAGCCTTGCAATTTAGCAGTTTACAAGGATAATGCAAGATTGATCACTTCTCTG 405  
QY 361 CTTACAGTGGAGCTGATATACAGAGGTTGATAGCGTGGGCTCACTGCCCTCCATATT 420  
DB 406 CTTACAGTGGAGCTGATATACAGAGGTTGATAGCGTGGGCTCACTGCCCTCCATATT 465  
QY 421 GCTACAACTAGTGGCACTGATAGGCTGCTGATGCTGTGCAACATGGAGCTAATGTC 480  
DB 466 GCTACAACTAGTGGCACTGATAGGCTGCTGATGCTGTGCAACATGGAGCTAATGTC 525  
QY 481 AATATTCAAGATCAGTTTTTTTTTCACTCCATTTGCAATTTGCAAGGCTATGGAATGAA 540  
DB 526 AATATTCAAGATCAGTTTTTTTTTCACTCCATTTGCAATTTGCAAGGCTATGGAATGAA 585  
QY 541 CAGTAACTCGCTCTCTTTTGAATTTGGTCTGATGTAATTAAGTGGTGAAGTCTGA 600  
DB 586 CAGTAACTCGCTCTCTTTTGAATTTGGTCTGATGTAATTAAGTGGTGAAGTCTGA 645  
QY 601 GATAGACCCCTCCACCTAGCATCTGCAAAAGGATTCTTGAATATTGCAAAACTTTGATG 660

DB 646 GATAGACCCCTCCACCTAGCATCTGCAAAAGGATTCTTGAATATTGCAAAACTTTGATG 705  
QY 661 GAAGAAGGAGCAAGCAAGATGTAATGCTCAAGTAATGAAGACCATGTCCCACTCCAT 720  
DB 706 GAAGAAGGAGCAAGCAAGATGTAATGCTCAAGTAATGAAGACCATGTCCCACTCCAT 765  
QY 721 TTCTGTTCTCGATTTGGACACCATGATAGTAAAGTATCTGCTCAAGTGAATTTGAA 780  
DB 766 TTCTGTTCTCGATTTGGACACCATGATAGTAAAGTATCTGCTCAAGTGAATTTGAA 825  
QY 781 GTTCAACCTCATGTTGTTAATATCTATGAGATACCCCTTTACACCTGGCATGTACAA 840  
DB 826 GTTCAACCTCATGTTGTTAATATCTATGAGATACCCCTTTACACCTGGCATGTACAA 885  
QY 841 GGCAAATTTGAATTTGCCAAGGAATCATCCAATATCAGGAACAGAAAGTCTGACTAAG 900  
DB 886 GGCAAATTTGAATTTGCCAAGGAATCATCCAATATCAGGAACAGAAAGTCTGACTAAG 945  
QY 901 GAAACATCTTCAGTGAAACAGCTTTTCATAGTGTGTTACTATGCGCAAGCATTTGAC 960  
DB 946 GAAACATCTTCAGTGAAACAGCTTTTCATAGTGTGTTACTATGCGCAAGCATTTGAC 1005  
QY 961 CTAGTCAAAATTTCTTCTGATCAGAAATGTCAATAACATCAACCAAGGAGGATGGG 1020  
DB 1006 CTAGTCAAAATTTCTTCTGATCAGAAATGTCAATAACATCAACCAAGGAGGATGGG 1065  
QY 1021 CACACTGGATTACACTCTGCTTGTACACAGGTCAATTCGCTGGTTTCACTTTACTG 1080  
DB 1066 CACACTGGATTACACTCTGCTTGTACACAGGTCAATTCGCTGGTTTCACTTTACTG 1125  
QY 1081 GATAATGGAGCTGATATGAATCTAGTGGCTTGATGCCAGAGCTCTAGTGGTGAATA 1140  
DB 1126 GATAATGGAGCTGATATGAATCTAGTGGCTTGATGCCAGAGCTCTAGTGGTGAATA 1185  
QY 1141 GATGACGACATGTTGATGTTGGCTTATGAAAGGCGATGATGCCATGTGCACATC 1200  
DB 1186 GATGACGACATGTTGATGTTGGCTTATGAAAGGCGATGATGCCATGTGCACATC 1245  
QY 1201 CTGAACCATTTAAGAGACCAAGATGAATTCGCTGTGAATTAATTTCTCAGCCTGA 1260  
DB 1246 CTGAACCATTTAAGAGACCAAGATGAATTCGCTGTGAATTAATTTCTCAGCCTGA 1305  
QY 1261 GGAGATGGCTCTATGTTGTTTCCATCAACCTTTGGGGAAGATTAAAGCATGCAAAA 1320  
DB 1306 GGAGATGGCTCTATGTTGTTTCCATCAACCTTTGGGGAAGATTAAAGCATGCAAAA 1365  
QY 1321 GAGAAGCAGATATTCTCTCTAAGAGCTGGATTGCCTTCACATTTCCATCTTCAGCTC 1380  
DB 1366 GAGAAGCAGATATTCTCTCTAAGAGCTGGATTGCCTTCACATTTCCATCTTCAGCTC 1425  
QY 1381 TCAGAAATTTGAGTTCCATGAGATTATGGCTCAGTTCTTTTGGGAAAGTATATAAGGA 1440  
DB 1426 TCAGAAATTTGAGTTCCATGAGATTATGGCTCAGTTCTTTTGGGAAAGTATATAAGGA 1485  
QY 1441 CGATGAGAAATTAATAGTGGCTATAAACCCTTATCGAGCCAACTACTGCTCCAAAG 1500  
DB 1486 CGATGAGAAATTAATAGTGGCTATAAACCCTTATCGAGCCAACTACTGCTCCAAAG 1545  
QY 1501 TCAGATGTGGATATGTTTGGCGAGAGGTGCCATTTCTGTCAGCTCAATCATCCCTGC 1560  
DB 1546 TCAGATGTGGATATGTTTGGCGAGAGGTGCCATTTCTGTCAGCTCAATCATCCCTGC 1605  
QY 1561 GTAATTCAGTTTGGTGGTCTTGTGTAATGATCCAGCAGTTTGCCATGTGCACTCAA 1620  
DB 1606 GTAATTCAGTTTGGTGGTCTTGTGTAATGATCCAGCAGTTTGCCATGTGCACTCAA 1665  
QY 1621 TACATATCAGGGGGTCTCTGTTCTCCCTCTCATGAGCAGAAAGGATTCTTGATTG 1680  
DB 1666 TACATATCAGGGGGTCTCTGTTCTCCCTCTCATGAGCAGAGGATTTCTTGATTG 1725  
QY 1681 CAGTCTAAATTAATTTATGTCAGTATGTTGCAAGGATGAGTACTTCAACAACCTG 1740

Db 1726 CAGTCTAAATTAATTATTGAGTAGATGTTGCCAAGGCGATGAGTACCTTCACAACCTG 1785  
QY 1741 ACACAGCCAAATATACATCGTGAATTCGAAAGTGCACAAATA 1780  
Db 1786 ACACAGCCAAATATACATCGTGAATTCGAAAGTGCATGTA 1825

## RESULT 10

ABT09088

ID ABT09088 standard; DNA; 668 BP.

XX AC

XX ABT09088;

XX DT 05-DEC-2002 (first entry)

XX DE Phase-1 Rat CT gene SEQ ID No 176.

XX KW Rat; toxicity study; rat toxic response gene; toxicological response;  
XX KW drug development; phase-1 rat CT gene; ds.  
XX OS Rattus sp.

XX PN WO200266682-A2.

XX PD 29-AUG-2002.

XX PF 29-JAN-2002; 2002WO-US002935.

XX PR 29-JAN-2001; 2001US-0264933P.

XX PR 26-JUL-2001; 2001US-0308161P.

XX PA (PHAS-) PHASE-1 MOLECULAR TOXICOLOGY INC.

XX PI Farris G, Hickien SH, Farr SB;

XX DR WPI; 2002-674961/72.

XX PT Evaluating the toxicity of an agent, useful in drug development or in  
XX PT determining toxicological responses to a new drug, by determining the  
XX PT expression of rat toxicologically relevant genes in the test animal in  
XX PT response to the test agent.  
XX PS Disclosure; Page 149; 388pp; English.

XX CC The invention relates to a method used for evaluating the toxicity of an  
XX CC agent comprising determining the expression of a rat toxic response  
XX CC gene(s) in the test animal in response to the agent. The method is useful  
XX CC in drug development, particularly for conducting toxicity studies and  
XX CC analysis before a new drug or compound is approved for human consumption  
XX CC or use. The method is also useful in determining toxicological responses  
XX CC to a new drug. This polynucleotide sequence represents a phase-1 rat CT  
XX CC gene of the invention

SQ Sequence 668 BP; 162 A; 186 C; 170 G; 149 T; 0 U; 1 Other;  
Query Match 18.1%; Score 466.4; DB 6; Length 668;  
Best Local Similarity 87.1%; Pred. No. 7e-124;  
Matches 512; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

QY 1702 GTAGATGTTGCCAAGGCGATGGAGTACCTTCACAACTGACAGCGCAATTTATCATCT 1761  
Db 77 GTAGAGCTTGCCAAAGGCGATGGAGTACCTTCACAACTGACCGCAATTCATACACCGC 136

QY 1762 GACTTGACAGTCAATATTTCTCTATGAGATGGCGATCGTGGTGCAGATTTT 1821  
Db 137 GACTTGACAGCCCAATTTCTCTCTATGAGATGGCGATCGTGGTGCAGATTTT 196

QY 1822 GGAGAATCAAGATTTCTACAGTCTCTCGATGAACAAATGACAAACAACTGGGAAC 1881  
Db 197 GGAGAATCAAGATTTCTGACGCTCCCTGGATGAAGACAAATGACAAAGCAGCGGAAC 256

QY 1882 CTCGTTGGATGGCTCTGAGTGTTCACGAGTGCATCGGTACACCATCAAGCAGAT 1941

Db 257 CTGCGTGGATGGCCCTGAGGTGTTTCACACAGTGCACGAGATACACCATCAAGCTGAT 316  
QY 1942 GTCTTCAAGCTATGCTGTGCTGTGGGAAATTCACCTGGGAAATTCATTTGCTCAT 2001  
Db 317 GTCTTCAAGTACTCCCTGTGCTGTGGGAGCTCCTCACTGGAGAAATTCATTTGCTCAT 376  
QY 2002 CTCAAGCCAGCGGCTGCGGCGAGCAGACATGGCTTACCACCATCAGACCTCCCATGGC 2061  
Db 377 CTCAAGCCAGCGGCTGCGGCGAGCAGACATGGCGTATACCCACATCAGACCGCCATGGC 436  
QY 2062 TATTCATTTCCAAAGCCCATATCATCTCTGCTGATACGAGGTGGAAGCATGCTCTGAA 2121  
Db 437 TATTCATTTCCAAAGCCCATATCATCTCTGCTGATACGAGGTGGAAGCATGCTCTGAA 496  
QY 2122 GGAAGACCCGAATTTTCTGAAGTTGTCATGAAGTTAGAAAGTGTCTCTGCAACATTCAG 2181  
Db 497 GGAGCACCAGATTTCTCTGAAGTCGTTAGCAAACTGGAGAGTGCCTATGCAATGTGAG 556  
QY 2182 CTGATGTTCTCTGATCAAGTAAAGTAAACAGCAGTGGTCTCTCTACCTTCTTCTTCTCAT 2241  
Db 557 CTCATGTTCTCCAGCATCAAGTAAACAGCAGTGGCTCTCTGTACCTTCTCTTCTTCCGAT 616  
QY 2242 TGCTGTGTGAACCGGGGAGGACCTGCGCGAGTCAATGTGGCAGCATTA 2289  
Db 617 TGCTGTGTGAACCGGGGAGGACCTGCGCGAGGAGTGCCTGCGGAGCCAGTGGCAGCCTTA 664

## RESULT 11

AAS33234

ID AAS33234 standard; cDNA; 633 BP.

XX AC AAS33234;

XX DT 04-DEC-2001 (first entry)

XX DE DNA encoding human secreted protein, Seq ID No 193.

XX KW Immunomodulatory; human immunodeficiency virus; HIV; anaemia; angina;  
XX KW rheumatoid arthritis; antiarteriosclerotic; cardiant; vascular;  
XX KW cerebroprotective; thrombolytic; antimicrobial; ophthalmological;  
XX KW cytostatic; Alzheimer's disease; Parkinson's disease; human; cancer;  
XX KW multiple sclerosis; cancer; hyperproliferative disorder; infection;  
XX KW Gaucher's disease; neurological disease; cerebrovascular disorder;  
XX KW thrombosis; wound healing; ss.

XX OS Homo sapiens.

XX PN WO200155326-A2.

XX PD 02-AUG-2001.

XX PF 17-JAN-2001; 2001WO-US001347.

XX PR 31-JAN-2000; 2000US-0179065P.

XX PR 04-FEB-2000; 2000US-0180628P.

XX PR 24-FEB-2000; 2000US-0184664P.

XX PR 02-MAR-2000; 2000US-0186350P.

XX PR 16-MAR-2000; 2000US-0189874P.

XX PR 17-MAR-2000; 2000US-019076P.

XX PR 18-APR-2000; 2000US-0198123P.

XX PR 07-JUN-2000; 2000US-0205515P.

XX PR 28-JUN-2000; 2000US-0209467P.

XX PR 30-JUN-2000; 2000US-0214886P.

XX PR 07-JUL-2000; 2000US-0215135P.

XX PR 07-JUL-2000; 2000US-0216647P.

XX PR 11-JUL-2000; 2000US-0216880P.

XX PR 11-JUL-2000; 2000US-0217487P.

XX PR 14-JUL-2000; 2000US-0217496P.

XX PR 26-JUL-2000; 2000US-0218290P.

XX PR 26-JUL-2000; 2000US-0220963P.

XX PR 14-AUG-2000; 2000US-0220964P.

XX PR 14-AUG-2000; 2000US-0224518P.

XX PR 14-AUG-2000; 2000US-0224519P.

PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226868P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 05-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 06-SEP-2000; 2000US-0231242P.  
PR 06-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 23-SEP-2000; 2000US-0234997P.  
PR 23-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239335P.  
PR 13-OCT-2000; 2000US-0239337P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-024617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.

PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246603P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249246P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 03-JAN-2001; 2001US-0259678P.  
XX  
XX  
(HUMA-) HUMAN GENOME SCI INC.  
Rosen CA, Barash SC, Ruben SM;  
WPI; 2001-451931/48.  
P-PSDB; AAU20525.  
New nucleic acids and polypeptides, useful for diagnosing, preventing or treating medical conditions.  
Claim 1; SEQ ID NO 193; 753pp; English.  
The invention relates to novel isolated nucleic acid molecules (I) encoding human secreted proteins (II). (I) and (II) are used to prevent, treat or ameliorate a medical condition in e.g. humans, mice, rabbits, goats, horses, cats, dogs, chickens or sheep. (I) and (II) may be used in the prevention, treatment and diagnosis of diseases associated with inappropriate expression of secreted proteins. (I) and complementary sequences may also be used as DNA probes in diagnostic assays (e.g. polymerase chain reactions (PCR)) to detect and quantitate the presence of similar nucleic acid sequences in samples, and so which patients may be in need of restorative therapy. (II) may also be used as antigens in the production of antibodies and in assays to identify modulators (agonists and antagonists) of the expression and activity of the secreted proteins. The anti-(II) antibodies and antagonists may also be used to down regulate expression and activity of (II). The anti-(II) antibodies may also be used as diagnostic agents for detecting the presence of (II) in samples (e.g. by enzyme linked immunosorbant assay (ELISA)). The disorders include for example: immune/autoimmune diseases (e.g. HIV





```

DB      463 AGATATTCCTCTCTAGGAGCTGGATTGCTTCACATTTCCATCTT 509
|||||
RESULT 13
AAS80722
ID      AAS80722 standard; cDNA; 1557 BP.
XX
AC      AAS80722;
XX
DT      13-FEB-2002 (first entry)
XX
DE      DNA encoding novel human diagnostic protein #16526.
XX
KW      Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW      food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS      Homo sapiens.
XX
PN      WO200175067-A2.
XX
PD      11-OCT-2001.
XX
PF      30-MAR-2001; 2001WO-US008631.
XX
PR      31-MAR-2000; 2000US-00540217.
PR      23-AUG-2000; 2000US-00649167.
XX
PA      (HYSE-) HYSEQ INC.
XX
XX      Dmanac RT, Liu C, Tang YT;
XX
XX      WPI; 2001-639362/73.
DR      P-PSDB; ABG16535.
XX
XX      New isolated polynucleotide and encoded polypeptides, useful in
PT      diagnostics, forensics, gene mapping, identification of mutations
PT      responsible for genetic disorders or other traits and to assess
PT      biodiversity.
XX
XX      Claim 1; SEQ ID NO 16526; 103pp; English.
XX
XX      The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC      sequences. (I) is useful as hybridisation probes, polymerase chain
CC      reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC      and in recombinant production of (II). The polynucleotides are also used
CC      in diagnostics as expressed sequence tags for identifying expressed
CC      genes. (I) is useful in gene therapy techniques to restore normal
CC      activity of (II) or to treat disease states involving (II). (II) is
CC      useful for generating antibodies against it, detecting or quantitating a
CC      polypeptide in tissue, as molecular weight markers and as a food
CC      supplement. (II) and its binding partners are useful in medical imaging
CC      of sites expressing (II). (I) and (II) are useful for treating disorders
CC      involving aberrant protein expression or biological activity. The
CC      polypeptide and polynucleotide sequences have applications in
CC      diagnostics, forensics, gene mapping, identification of mutations
CC      and to produce other types of data and products dependent on DNA and
CC      amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC      coding sequences of the invention. Note: The sequence data for this
CC      patent did not appear in the printed specification, but was obtained in
CC      electronic format directly from WIPO at
CC      ftp.wipo.int/pub/published_pct_sequences
XX
XX      Sequence 1557 BP; 417 A; 400 C; 459 G; 281 T; 0 U; 0 Other;
SQ
Query Match      5.4%; Score 135.6; DB 5; Length 1557;
Best Local Similarity 68.4%; Pred. No. 4.8e-25;
Matches 245; Conservative 0; Mismatches 59; Indels 54; Gaps 2;

DB      40 GATGAATCGAAGAAAAAGTCAAGTGAATCATATGTTATCAATAGAAAGATTAGAAGAT 99
178 GATGAATCGAAGAAAAAGTCAAGTGAATCATATGTTATCAATAGAAAGATTAGAAGAT 237
|||||
100 GACCTGCAGATCAAGGAAAAAGAACTGACAGAACTAAGGAATATATTTGGCTCTCATGAA 159
|||||
238 GACCTGCAGATCAAGGAAAAAGAACTGACAGAACTAAGGAATATATTTGGTGGTTTGTG 297
|||||
160 GCCTT-----CAGTAA 171
|||||
298 ATCTTGTGCTTTCAGGAGTGAAGCTGCAGGCTTCATGGTGAAGTGTACAGCTCATAAA 357
|||||
172 GTCAATTTAAATTACCGCACTGAAATGGCTGTCTCTACTTCA-----TTTATGTTTC 225
|||||
358 GGCAGTCCGACCCCAAGAGTGAAGAGCAGCAAGATTATTGCAAAAAGCGAAGACAA 417
|||||
226 ATTTGTGGAGGCAAGAAATACATATTCGAACTCTTATGTTGAAAGGGCTCGGCCATCT 285
|||||
418 AGCTCCCAAGCAAGAAATCACATATTCGAACTCTTATGTTGAAAGGGCTCGGCCATCT 477
|||||
286 CGACTGACAAAGAAATGGAATTTACAGCTTGCAGCTTTCAGAGGATTAATGCAG 343
|||||
478 CGACTGACAAAGAAATGGAATTTACAGCTTGCAGCTTTCAGAGGATTAATGCAG 535
|||||

RESULT 14
AAC48526
ID      AAC48526 standard; DNA; 2283 BP.
XX
AC      AAC48526;
XX
DT      18-OCT-2000 (first entry)
XX
DE      Arabidopsis thaliana DNA fragment SEQ ID NO: 57797.
XX
KW      Hybridisation assay; genetic mapping; gene expression control;
KW      protein identification; signal transduction pathway; metabolic pathway;
KW      promoter; termination sequence; ss.
XX
XX      Arabidopsis thaliana.
OS
XX
XX      EP1033405-A2.
PN
XX
PD      06-SEP-2000.
XX
XX      25-FEB-2000; 2000EP-00301439.
PF
XX
XX      25-FEB-1999; 99US-0121825P.
XX      05-MAR-1999; 99US-0123180P.
XX      09-MAR-1999; 99US-0123548P.
XX      23-MAR-1999; 99US-0125788P.
XX      25-MAR-1999; 99US-0126264P.
XX      29-MAR-1999; 99US-0126785P.
XX      01-APR-1999; 99US-0127462P.
XX      06-APR-1999; 99US-0128234P.
XX      08-APR-1999; 99US-0128714P.
XX      16-APR-1999; 99US-0129845P.
XX      19-APR-1999; 99US-0130077P.
XX      21-APR-1999; 99US-0130449P.
XX      23-APR-1999; 99US-0130510P.
XX      23-APR-1999; 99US-0130891P.
XX      28-APR-1999; 99US-0131449P.
XX      30-APR-1999; 99US-0132048P.
XX      30-APR-1999; 99US-0132407P.
XX      04-MAY-1999; 99US-0132484P.
XX      05-MAY-1999; 99US-0132485P.
XX      06-MAY-1999; 99US-0132486P.
XX      07-MAY-1999; 99US-0132487P.
XX      11-MAY-1999; 99US-0132863P.
XX      14-MAY-1999; 99US-0134218P.
XX      14-MAY-1999; 99US-0134219P.
XX      14-MAY-1999; 99US-0134221P.
XX      14-MAY-1999; 99US-0134370P.
XX      18-MAY-1999; 99US-0134768P.
XX      19-MAY-1999; 99US-0134941P.

```

PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 24-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 08-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 23-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 20-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 06-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 03-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151103P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158322P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159333P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.



Db 1504 ATACCAAAATGCTCATGCAAACTTCTGAGCTTCTTCAGAAATGTTGGCAACAGGAG 1563  
Qy 2116 CCTGAGGAGAGACCGGAATTTCTGAAGT 2144  
Db 1564 CCTGCTGAAAGACCAGATTTCTCTGAAT 1592

Search completed: September 6, 2004, 01:05:54  
Job time : 636.459 secs

This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 2, 2004, 14:07:41 ; Search time 136 Seconds  
(without alignments)  
1734.759 Million cell updates/sec

Title: US-10-626-173-2

Perfect score: 4390

Sequence: 1 MGNYKSRPTQCTDEWKKV.....PMSSMHFHSRNSSSFFEDSS 835

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Genesep\_29Jan04:\*

- 1: Genesep1980s:\*
- 2: Genesep1980s:\*
- 3: Genesep2000s:\*
- 4: Genesep2001s:\*
- 5: Genesep2002s:\*
- 6: Genesep2003as:\*
- 7: Genesep2003bs:\*
- 8: Genesep2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4390	100.0	835	3 AAB01470	Aab01470 Human CAR
2	4390	100.0	835	4 AAB65674	Aab65674 Novel pro
3	4390	100.0	835	6 ABR41886	AbR41886 Human car
4	4311	98.2	928	4 ABG16533	Abg16533 Novel hum
5	4079	92.9	835	3 AAB01474	AaB01474 Rat CAR
6	4079	92.9	835	6 ABR41887	AbR41887 Rat cardi
7	3102	70.7	603	6 AAE37962	Aae37962 Human kin
8	603	13.7	148	4 AAU20525	Aau20525 Human sec
9	603	13.7	148	4 AAU20648	Aau20648 Human sec
10	493	11.2	589	3 AAG45984	Aag45984 Arabidops
11	493	11.2	732	3 AAG45983	Aag45983 Arabidops
12	493	10.5	1881	7 AAG45982	Aag45982 Arabidops
13	460.5	10.3	1536	6 ABU11523	Abu11523 Human pro
14	454	10.3	1536	6 ABU11523	Abu11523 Human pro
15	446	10.2	1549	4 ABB67412	Abb67412 Drosophil
16	446	10.2	1549	4 ABB67412	Abb67412 Drosophil
17	442.5	10.1	1056	6 AAE33684	Aae33684 Human str
18	440.5	10.0	1053	4 AAU28174	Aau28174 Novel hum
19	440	10.0	705	4 AAU20496	Aau20496 Human sec
20	438	10.0	412	3 AAG22170	Aag22170 Arabidops
21	437.5	10.0	369	3 AAG22172	Aag22172 Arabidops
22	437.5	10.0	374	3 AAG22171	Aag22171 Arabidops
23	435	9.9	489	4 ABUS3032	AbuS3032 Cell stru
24	435	9.9	1839	7 ADD27862	AdD27862 Human bra
25	435	9.9	1872	4 AAM79160	Aam79160 Human pro

## ALIGNMENTS

### RESULT 1

AAB01470  
ID AAB01470 standard; protein; 835 AA.

XX AAB01470;

XX AC

XX 20-OCT-2000 (first entry)

XX DT Human CAR (Cardiac related Ankyrin-Repeat Protein Kinase).

XX DE

XX KW Cardiac related ankyrin repeat protein kinase; CAR; cytoskeleton;

XX KW cardiac cell growth factor receptor; cell differentiation; modulator;

XX KW regulator; detection; cellular proliferation; cardiovascular disorder;

XX KW heart failure; hypertension; cancer; sarcoma.

XX OS Homo sapiens.

XX XX

XX WO200034330-A1.

XX XX

XX PD 15-JUN-2000.

XX XX

XX PF 10-DEC-1999; 99WO-US029465.

XX PR 11-DEC-1998; 98US-0111938P.

XX PR 14-APR-1999; 99US-00291839.

XX PA (MILL-) MILLENNIUM PHARM INC.

XX XX

XX PI Raju J;

XX XX

XX WPI; 2000-431275/37.

XX DR N-PSDB; AAA47606, AAA47607.

XX XX

XX PT New polynucleotide encoding cardiac-related ankyrin-repeat protein

XX PT kinase, useful for treating disorders such as cardiovascular disorders,

XX PT e.g. heart failure and cell differentiation disorders, e.g. cancer.

XX XX

XX PS Claim 2; Fig 1; 161pp; English.

XX CC CARP polypeptides are regulators of signal transmission from cellular

XX CC receptors (e.g. cardiac cell growth factor receptor) and cytoskeleton

XX CC function. They also act as modulators of cellular differentiation and

XX CC cell death. Nucleic acids, or their fragments encoding CARP polypeptides

XX CC are useful for detecting CARP nucleic acids especially mRNA, in a sample.

XX CC CARP is useful for treating disorders associated with upregulation or

XX CC downregulation of cellular proliferation such as, cardiovascular

XX CC disorders (heart failure and hypertension) and disorders associated with

XX CC cell differentiation such as cancer and sarcoma

XX CC

XX CC

XX	Sequence	835 AA;
XX	Human; mouse; protein kinase; antiarthritic; antisclerotic; osteopathic; immunosuppressive; cardiac; renal; antiinflammatory; antisthmatic; dermatological; antidiabetic; antifertility; gene therapy; vaccine; immune disorder; cardiovascular disease; neurodegenerative disease; cancer; autoimmune disorder; stroke; inflammatory bowel disease; inflammatory pelvic disease; multiple sclerosis; psoriasis.	
XX	Homo sapiens.	
OS	WO200073469-A2.	
PN	07-DEC-2000.	
XX	26-MAY-2000; 2000WO-US014842.	
XX	28-MAY-1999; 99US-0136503P.	
PR	(SUGEN-) SUGEN INC.	
XX	Plowman GD, Martinez R, Whyte D, Sudersanam S;	
PI	WPI; 2001-032161/04.	
DR	N-PSDB; AAF44702.	
XX	Nucleic acids encoding kinase polypeptides, useful for diagnosing and treating immune-related diseases and disorders, cardiovascular disease, neurodegenerative diseases and/or cancers.	
PT	Claim 10; Fig 1; 310pp; English.	
XX	The present sequence is a novel protein kinase. The novel protein kinases and the nucleic acids that encode them may be used in the treatment and diagnosis of diseases associated with inappropriate kinase expression such as immune-related diseases and disorders, cardiovascular disease, neurodegenerative diseases and/or cancers. The nucleic acids and complementary sequences may also be used as DNA probes in diagnostic assays. The kinase polypeptides may be used as antigens in the production of antibodies of kinase expression and activity. Anti-kinase antibodies and kinase antagonists may also be used to down regulate kinase expression and activity. Diseases related to kinase expression and activity include rheumatoid arthritis, atherosclerosis, autoimmune disorders, complications of organ transplantation, myocardial infarction, immune disorders, cardiomyopathies, strokes, renal failure, oxidative stress related disorders, chronic inflammatory bowel disease, chronic inflammatory pelvic disease, multiple sclerosis, asthma, osteoarthritis, psoriasis, rhinitis, autoimmunity, diabetes, cancers and reproductive disorders	
XX	Sequence 835 AA;	
XX	Query Match	100.0%; Score 4390; DB 4; Length 835;
XX	Best Local Similarity	100.0%; Pred. No. 0;
XX	Matches	835; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 MGNYSRPTQCTDEWKKVSSVYITIERLEDDLOIKKELTELNRNIFGSDFAFSKVN	60
DB	1 MGNYSRPTQCTDEWKKVSSVYITIERLEDDLOIKKELTELNRNIFGSDFAFSKVN	60
QY	61 NYRTENGLSLHLCCTCGGKSHIRTLMLKGLRPSRLTRNGFTALHLAVYKDAELITSL	120
DB	61 NYRTENGLSLHLCCTCGGKSHIRTLMLKGLRPSRLTRNGFTALHLAVYKDAELITSL	120
QY	121 LHSGADIQQVGYGGLTALHIATIAHLEAADVLLQGANVNIQDAVFTPLHIAAYGHE	180
DB	121 LHSGADIQQVGYGGLTALHIATIAHLEAADVLLQGANVNIQDAVFTPLHIAAYGHE	180
QY	181 QVTRLLKFGADVNVSGEVDGDRPLHLASAKGFLNAKLMEEGSKADVNAQDNEDHVP	240
DB	181 QVTRLLKFGADVNVSGEVDGDRPLHLASAKGFLNAKLMEEGSKADVNAQDNEDHVP	240
QY	241 FCSRFGHHDIVKYLQSDLEVPVNVNIYGDTPHLACVNGKFEVAKI	300
DB	241 FCSRFGHHDIVKYLQSDLEVPVNVNIYGDTPHLACVNGKFEVAKI	300
QY	301 ENIFSETAHSACTYCKSDIAVKFLLDQNVININHQGRDGTGLHSACVGHILVQFLL	360
DB	301 ENIFSETAHSACTYCKSDIAVKFLLDQNVININHQGRDGTGLHSACVGHILVQFLL	360
QY	361 DNGADNVLVACDPSRSGSEKEDQTCIMWAVEYKGDHAI	420
DB	361 DNGADNVLVACDPSRSGSEKEDQTCIMWAVEYKGDHAI	420
QY	421 GDGSYVSPSPGKIKSMYKADILLRAGLPSHFHQLSEIFPHEIIGSGSGKVKYG	480
DB	421 GDGSYVSPSPGKIKSMYKADILLRAGLPSHFHQLSEIFPHEIIGSGSGKVKYG	480
QY	481 RCNKIVALKRYRANTYCSKSDVDMFCREVSILCOLNHPCVIQFVGACLNPPSOFAIVTQ	540
DB	481 RCNKIVALKRYRANTYCSKSDVDMFCREVSILCOLNHPCVIQFVGACLNPPSOFAIVTQ	540
QY	541 YISGGSFLSLLHEQKRLDLSKLIIVADVAKGMEYLNLTQPIIHRDLNSHNILLYEDG	600
DB	541 YISGGSFLSLLHEQKRLDLSKLIIVADVAKGMEYLNLTQPIIHRDLNSHNILLYEDG	600
QY	601 HAVVADFGESRFLQSDLDENMTKQGNLRWMAPEVFTQCTRYT	660
DB	601 HAVVADFGESRFLQSDLDENMTKQGNLRWMAPEVFTQCTRYT	660
QY	661 GEIPFAHLKPAADAAADMAHYHIRPIGYSIPKPISSLLIRGNWACPEGRPFSEVVMKLE	720
DB	661 GEIPFAHLKPAADAAADMAHYHIRPIGYSIPKPISSLLIRGNWACPEGRPFSEVVMKLE	720
QY	721 ECLCNTELMPASSNSSLSPSSSDCLVNRGGPRSHVAALRSRFELEYALNARSYAA	780
DB	721 ECLCNTELMPASSNSSLSPSSSDCLVNRGGPRSHVAALRSRFELEYALNARSYAA	780
QY	781 LSQSAQYSSQGLSLEEMKRSIQYTPIDKYGVSDPMSMSEHFCRNSSSPEDSS	835
DB	781 LSQSAQYSSQGLSLEEMKRSIQYTPIDKYGVSDPMSMSEHFCRNSSSPEDSS	835
XX	RESULT 2	
XX	AAB65674	
XX	ID AAB65674 standard; protein; 835 AA.	
XX	AC AAB65674;	
XX	DT 27-MAR-2001 (first entry)	
XX	DE Novel protein kinase, SEQ ID NO: 202.	

```
Db 241 FCSRGHHDIYKLLQSLDEVPQVNVNIGDTPHLACYNKGKFEVAKETIIQISGTESLTK 300
Qy 301 ENIFSETAFHSACTYKSKIDLAKFLLDONVININHGDRGHTGLHSACVHGHIRLVOQLL 360
Db 301 ENIFSETAFHSACTYKSKIDLAKFLLDONVININHGDRGHTGLHSACVHGHIRLVOQLL 360
Qy 361 DNGADNMLVACDPSRSGGKDEQTCMLWAYEXGHDIAIVTLLKHRYRPODELPCNEYSOPG 420
Db 361 DNGADNMLVACDPSRSGGKDEQTCMLWAYEXGHDIAIVTLLKHRYRPODELPCNEYSOPG 420
Qy 421 GDGSYVSPSPGLKIKSMYKEXKADILLRAGLPSFHLQLSIEFHEIIGSGSPGKYKG 480
Db 421 GDGSYVSPSPGLKIKSMYKEXKADILLRAGLPSFHLQLSIEFHEIIGSGSPGKYKG 480
Qy 481 RCNKIVAIRYRANTYCKSDVDMFCREVSILCOLNHPCVIOFVGACLNPSQFAIVTQ 540
Db 481 RCNKIVAIRYRANTYCKSDVDMFCREVSILCOLNHPCVIOFVGACLNPSQFAIVTQ 540
Qy 541 YISGGSLSLHQRKILDLQSKLIITAVDAKGMAYLHNLTOPIIHRDLNHNILLYEDG 600
Db 541 YISGGSLSLHQRKILDLQSKLIITAVDAKGMAYLHNLTOPIIHRDLNHNILLYEDG 600
Qy 601 HAVVADFGSRFLQSLDENMTKQPNLWMAPEVPTQCTRTYIKADYFVALCLWEILT 660
Db 601 HAVVADFGSRFLQSLDENMTKQPNLWMAPEVPTQCTRTYIKADYFVALCLWEILT 660
Qy 661 GEIPFAHLKFAAAADWYHRIPIGYSIPKPISSLIRGNWACPEGRPFSEVVMKLE 720
Db 661 GEIPFAHLKFAAAADWYHRIPIGYSIPKPISSLIRGNWACPEGRPFSEVVMKLE 720
Qy 721 ECLCNTELSPASSNSGSLSPSSSDCLVNRGGGRSHVAALRRFLEVALNARSVAA 780
Db 721 ECLCNTELSPASSNSGSLSPSSSDCLVNRGGGRSHVAALRRFLEVALNARSVAA 780
Qy 781 LSQSAQYSSQGLSLEEMKRSIQYTPIDYGVVSDPMSSMHPHSCRNSSPEDSS 835
Db 781 LSQSAQYSSQGLSLEEMKRSIQYTPIDYGVVSDPMSSMHPHSCRNSSPEDSS 835

RESULT 3
ABR41886
ID ABR41886 standard; protein; 835 AA.
XX
XX
AC ABR41886;
XX
XX
DT 11-AUG-2003 (first entry)
DE
XX
XX
XX Human cardiac-related ankyrin-repeat protein kinase (CARK).
XX Human; CARK; cardiac-related ankyrin-repeat protein kinase; enzyme;
XX Cardiant; hypertensive; cytostatic.
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FH 2..7
FT Modified-site /note= "N-myristoylation motif"
FT Modified-site 11..14
FT Modified-site 18..21
FT Modified-site /note= "potential casein kinase II phosphorylation site"
FT Modified-site /note= "potential cAMP/cGMP dependent protein kinase phosphorylation site"
FT Modified-site 50..55
FT Modified-site /note= "N-myristoylation motif"
FT Domain 66..99
FT Modified-site /note= "Ankyrin repeat domain"
FT Modified-site 78..83
FT Modified-site /note= "N-myristoylation motif"
FT Modified-site 91..96
FT Domain 100..132
FT Modified-site /note= "Ankyrin repeat domain"
FT Modified-site 123..126

/note= "potential casein kinase II phosphorylation site"
133..165
/note= "ankyrin repeat domain"
133..138
/note= "N-myristoylation motif"
157..162
/note= "N-myristoylation motif"
168..198
/note= "ankyrin repeat domain"
190..195
/note= "N-myristoylation motif"
194..197
/note= "potential N-glycosylation site"
199..233
/note= "ankyrin repeat domain"
208..210
/note= "potential protein kinase C phosphorylation site"
224..227
/note= "potential casein kinase II phosphorylation site"
234..268
/note= "ankyrin repeat domain"
257..260
/note= "potential casein kinase II phosphorylation site"
269..302
/note= "ankyrin repeat domain"
293..296
/note= "potential casein kinase II phosphorylation site"
306..338
/note= "ankyrin repeat domain"
339..371
/note= "ankyrin repeat domain"
343..348
/note= "N-myristoylation motif"
363..368
/note= "N-myristoylation motif"
376..379
/note= "potential casein kinase II phosphorylation site"
411..415
/note= "LXCXE motif"
420..425
/note= "N-myristoylation motif"
461..464
/note= "potential casein kinase II phosphorylation site"
483..716
/note= "protein kinase domain"
489..477
/note= "protein kinase ATP binding region"
499..502
/note= "potential casein kinase II phosphorylation site"
526..531
/note= "N-myristoylation motif"
545..550
/note= "N-myristoylation motif"
579..582
/note= "potential N-glycosylation site"
615..618
/note= "potential casein kinase II phosphorylation site"
620..623
/note= "potential N-glycosylation site"
643..645
/note= "potential protein kinase C phosphorylation site"
735..738
/note= "potential N-glycosylation site"
744..747
/note= "potential casein kinase II phosphorylation site"
784..771
/note= "potential tyrosine kinase phosphorylation site"
765..768
/note= "potential casein kinase II phosphorylation site"
794..797
/note= "potential casein kinase II phosphorylation site"
805..808
/note= "potential casein kinase II phosphorylation site"
```

FT Modified-site 824..826 /note= "potential protein kinase C phosphorylation site"  
 FT Modified-site 827..830 /note= "potential N-glycosylation site"  
 FT Modified-site 829..832 /note= "potential casein kinase II phosphorylation site"  
 FT Modified-site 876..791 /note= "N-myristoylation motif"  
 FT WO2003020912-A2.  
 PN 13-MAR-2003.  
 PD 04-SEP-2002; 2002WO-US028300.  
 PP 05-SEP-2001; 2001US-00947199.  
 XX (MILL-) MILLENNIUM PHARM INC.  
 PA Raju J;  
 PI WPI; 2003-290188/28.  
 DR N-PSDB; ACC48580.  
 DR Novel isolated cardiac-related ankyrin-repeat protein kinase polypeptide,  
 PT useful for treating cellular growth related disorders which include  
 PT cardiovascular disorders and proliferative and/or differentiative  
 PT disorders.  
 XX Claim 14; Fig 1B; 158pp; English.

The present sequence is the protein sequence of a novel 96 kDa human cardiac-related ankyrin-repeat protein kinase, designated CARX, which plays a role in signalling pathways associated with cardiac cell growth or differentiation. CARX is a dual specificity protein kinase having specificity for both serine/threonine and tyrosine residues. CARX gene expression is increased in ischaemic heart tissue samples, suggesting a role in the regulation of cardiac cell growth and/or differentiation and the pathogenesis of cardiovascular disorders, e.g. congestive heart failure and cardiac hypertrophy. The invention provides CARX proteins, fusion proteins, antigenic peptides and anti-CARX antibodies, and also CARX nucleic acids, antisense molecules, recombinant expression vectors, host cells and transgenic animals. These are useful as modulating agents for regulating a variety of cellular processes, e.g. cardiac cellular process, for modulating the phosphorylation state of a CARX molecule or one or more proteins involved in cellular growth or differentiation, for modulating cell behaviour or as targets and therapeutic agents controlling cardiac cell proliferation, differentiation, hypertrophy and migration for modulating intra- or inter-cellular signalling and/or gene transcription, for modulating cell proliferation, growth, differentiation, survival and/or migration, for regulating transmission of signals from cellular receptors, for modulating entry of cells, e.g. cardiac precursor cells, into mitosis, or for regulating cytoskeletal function. The proteins and nucleic acids are useful for treating cellular growth related disorders which include cardiovascular disorders (such as heart failure, hypertension), and proliferative and/or differentiative disorders (such as cancer). They are also useful in screening assays, detection assays (e.g. chromosomal mapping, tissue typing, forensic biology), predictive medicine (e.g., diagnostic assays, prognostic assays, monitoring clinical trials) and pharmacogenomics, and in methods of treatment. A CARX polypeptide can be used to generate a specific antibody, to screen for naturally-occurring CARX substrates, and to screen for drugs or compounds which modulate CARX activity. It is also useful as a bait protein in a yeast two-hybrid or three-hybrid assay and for identifying other proteins which bind to, or interact with, CARX or which are involved in CARX activity

Sequence 835 AA;

Query Match 100.0%; Score 4390; DB 6; Length 835;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 835; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MGNYSRPTCTCTDEWKKVSESYVITIERLEDDLOIKEKELTELNRNIFGSDFAFSKYNL	60
DB	1	MGNYSRPTCTCTDEWKKVSESYVITIERLEDDLOIKEKELTELNRNIFGSDFAFSKYNL	60
QY	61	NYRTENGLSLHLHCCICGGKKSHIRMLKGRPSRLTRNGFTALHLAVYKDNABELITSL	120
DB	61	NYRTENGLSLHLHCCICGGKKSHIRMLKGRPSRLTRNGFTALHLAVYKDNABELITSL	120
QY	121	LHSGADIQQVGYGGLTALHATATAGLEADVLLQHGANNVIOQAVFFTPPLHIAAYGHE	180
DB	121	LHSGADIQQVGYGGLTALHATATAGLEADVLLQHGANNVIOQAVFFTPPLHIAAYGHE	180
QY	181	QVTRLLKFKGADVNVSGVGDRLPLHLASAKGFLNIAKLMEERGSKADVNAQDNEDHVLH	240
DB	181	QVTRLLKFKGADVNVSGVGDRLPLHLASAKGFLNIAKLMEERGSKADVNAQDNEDHVLH	240
QY	241	FCSRFGHHDIVKYLLOSLDELEVQPHVNIYGDTPHLACVNGKFEVAKELIQTSGTESLTK	300
DB	241	FCSRFGHHDIVKYLLOSLDELEVQPHVNIYGDTPHLACVNGKFEVAKELIQTSGTESLTK	300
QY	301	ENTFSETAFHSACTYGKSIDLVKFLDQNVININHOGRDHTGLHSACVHGHIRLVQFLL	360
DB	301	ENTFSETAFHSACTYGKSIDLVKFLDQNVININHOGRDHTGLHSACVHGHIRLVQFLL	360
QY	361	DNGADMLVACDPSRSGEKDEQTCMLMAYEKGHDAIVTLKHYYKRPQDELPCNYSQFG	420
DB	361	DNGADMLVACDPSRSGEKDEQTCMLMAYEKGHDAIVTLKHYYKRPQDELPCNYSQFG	420
QY	421	GDGSYVSPPLGKIKSMTEKADILLRAGLPSHFHLQLSIEFHEIIGSGFGKYVG	480
DB	421	GDGSYVSPPLGKIKSMTEKADILLRAGLPSHFHLQLSIEFHEIIGSGFGKYVG	480
QY	481	RCRNKIVAIKRYRANTYCSKSDVDMFCREVSILOLNHPFCVIOFVGACLNDSQFAIVTQ	540
DB	481	RCRNKIVAIKRYRANTYCSKSDVDMFCREVSILOLNHPFCVIOFVGACLNDSQFAIVTQ	540
QY	541	YISGGSLSLHLEQKRLDLSKLITAVDVAKGMEYLNLTQPIIHRDLNSHNILLYEDG	600
DB	541	YISGGSLSLHLEQKRLDLSKLITAVDVAKGMEYLNLTQPIIHRDLNSHNILLYEDG	600
QY	601	HAVVADPGESRFQSLDEDNMTKQPGNLRMAPEVFTQCTRYTIKADVFSYALCLWEILT	660
DB	601	HAVVADPGESRFQSLDEDNMTKQPGNLRMAPEVFTQCTRYTIKADVFSYALCLWEILT	660
QY	661	GEIPFAHLKPAALAAAMAYHHIRPPIGYIPKPISSLLIRGNVACPEGEPESEVVMKLE	720
DB	661	GEIPFAHLKPAALAAAMAYHHIRPPIGYIPKPISSLLIRGNVACPEGEPESEVVMKLE	720
QY	721	ECLCNIELMSPASSNSGSLSPSSSDCLVNRGGPGRSHVAALRSRFELEYALNARSYAA	780
DB	721	ECLCNIELMSPASSNSGSLSPSSSDCLVNRGGPGRSHVAALRSRFELEYALNARSYAA	780
QY	781	LSQSAGQYSSQGLSLEEMKRSLOQYTPIDKYGVYVDPMSMHFHSCRNSSSPFDDSS	835
DB	781	LSQSAGQYSSQGLSLEEMKRSLOQYTPIDKYGVYVDPMSMHFHSCRNSSSPFDDSS	835
RESULT 4			
ABG16533		ABG16533 standard; protein; 928 AA.	
XX		ABG16533;	
AC		ABG16533;	
DT		18-FEB-2002 (first entry)	
XX		Novel human diagnostic protein #16524.	
KW		Human; chromosome mapping; gene mapping; gene therapy; forensic;	
KW		food supplement; medical imaging; diagnostic; genetic disorder.	
OS		Homo sapiens.	
XX			
PN		WO200175067-A2.	

XX 11-OCT-2001.  
XX 30-MAR-2001; 2001WO-US008631.  
XX 31-MAR-2000; 2000US-00540217.  
XX 23-AUG-2000; 2000US-00649167.  
XX (HYSE-) HYSEQ INC.  
XX Dmanac RT, Liu C, Tang YT;  
XX WPI; 2001-639362/73.  
XX N-PSDB; AAS80720.  
XX New isolated polynucleotide and encoded polypeptides, useful in  
XX diagnostics, forensics, gene mapping, identification of mutations  
XX responsible for genetic disorders or other traits and to assess  
XX biodiversity.  
XX Claim 20; SEQ ID NO 46892; 103pp; English.  
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)  
XX sequences. (I) is useful as hybridisation probes, polymerase chain  
XX reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
XX and in recombinant production of (II). The polynucleotides are also used  
XX in diagnostics as expressed sequence tags for identifying expressed  
XX genes. (I) is useful in gene therapy techniques to restore normal  
XX activity of (II) or to treat disease states involving (II). (II) is  
XX useful for generating antibodies against it, detecting or quantitating a  
XX polypeptide in tissue, as molecular weight markers and as a food  
XX supplement. (II) and its binding partners are useful in medical imaging  
XX of sites expressing (II). (I) and (II) are useful for treating disorders  
XX involving aberrant protein expression or biological activity. The  
XX polypeptide and polynucleotide sequences have applications in  
XX diagnostics, forensics, gene mapping, identification of mutations  
XX responsible for genetic disorders or other traits to assess biodiversity  
XX and to produce other types of data and products dependent on DNA and  
XX amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic  
XX amino acid sequences of the invention. Note: The sequence data for this  
XX patent did not appear in the printed specification, but was obtained in  
XX electronic format directly from WIPO at  
XX ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 928 AA;  
Query Match 98.2%; Score 4311; DB 4; Length 928;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 821; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
13 TDWKKKVSSEYVITIERLEDDLOIKELTTELNRNIFGDEAFSKVNLNRYTENGSLH 72  
106 SDEWKKKVSSEYVITIERLEDDLOIKELTTELNRNIFGDEAFSKVNLNRYTENGSLH 165  
73 LCCICGGKSHITLMKGLRPSLFRNGFTALHLAVYKDNALITSLHSGADIQQVGY 132  
166 LCCICGGKSHITLMKGLRPSLFRNGFTALHLAVYKDNALITSLHSGADIQQVGY 225  
133 GGLTALHIAIAGHLEAADVLLQHGANNVIOQAVFPTPLHIAAYYGHEQVTRLLKFGAD 192  
226 GGLTALHIAIAGHLEAADVLLQHGANNVIOQAVFPTPLHIAAYYGHEQVTRLLKFGAD 285  
193 VNVSGEVDRPLHLASAKGFLNTAKLLMBEGSKADVNAQDNEDHVPFLHCFRFGHHDIVK 252  
286 VNVSGEVDRPLHLASAKGFLNTAKLLMBEGSKADVNAQDNEDHVPFLHCFRFGHHDIVK 345  
253 YLLQSDLEVOPHVNVYIGDTPHLACVNGKFEVAKELIIGSTESLTKENIFSETAFHSA 312  
346 YLLQSDLEVOPHVNVYIGDTPHLACVNGKFEVAKELIIGSTESLTKENIFSETAFHSA 405  
313 CTYGKSIDLVKFLDDQNVNINHQGRDGTGLHSAHYGHIRLVQFLDNGADMLNVACD 372  
406 CTYGKSIDLVKFLDDQNVNINHQGRDGTGLHSAHYGHIRLVQFLDNGADMLNVACD 465

QY 373 PFRSSGEKDEQTCMLWAYEKGHDAIVTLLKHKKRQDELPCNEYSQPGDGSYVSPPL 432  
DB 466 PFRSSGEKDEQTCMLWAYEKGHDAIVTLLKHKKRQDELPCNEYSQPGDGSYVSPPL 525  
QY 433 GKI KSMTEKADILLRAGLPSHFHLQSEIEFHEIIGSGSFGKVKYKRCRNKIVAIKRY 492  
DB 526 GKI KSMTEKADILLRAGLPSHFHLQSEIEFHEIIGSGSFGKVKYKRCRNKIVAIKRY 585  
QY 493 RANTYCSKSDVDMFCREVSILCOLNHPCVIOFVGACLNDRPSQFAIVTOYISGGSLSFLH 552  
DB 586 RANTYCSKSDVDMFCREVSILCOLNHPCVIOFVGACLNDRPSQFAIVTOYISGGSLSFLH 645  
QY 553 EOKRIIDLOSLLIIADVAKGMEYLHNLTQPTIHRDLNSHNIILYEDGHVAVVDFGESRF 612  
DB 646 EOKRIIDLOSLLIIADVAKGMEYLHNLTQPTIHRDLNSHNIILYEDGHVAVVDFGESRF 705  
QY 613 LQSLDEDNMTKQGNLRWMAPEVFTQCTRYTIKADVPFYSYALCLMEILLTGEIPFAHLKPA 672  
DB 706 LQSLDEDNMTKQGNLRWMAPEVFTQCTRYTIKADVPFYSYALCLMEILLTGEIPFAHLKPA 765  
QY 673 AADNAYHHRPIGYSIPKPISSILLIRGNACPSRPFSEVVMKLEECNIELMSPA 732  
DB 766 AADNAYHHRPIGYSIPKPISSILLIRGNACPSRPFSEVVMKLEECNIELMSPA 825  
QY 733 SSNSGSLSPSSSSDCLVNRGGPGRSHVAALRSRFELEYALNARSYAALSQAGYSQ 792  
DB 826 SSNSGSLSPSSSSDCLVNRGGPGRSHVAALRSRFELEYALNARSYAALSQAGYSQ 885  
QY 793 LSLEEMKRSLOQTPDKYGVSDPMSSMHFSCRNSSSFEDSS 835  
DB 886 LSLEEMKRSLOQTPDKYGVSDPMSSMHFSCRNSSSFEDSS 928  
RESULT 5  
AAB01474  
ID AAB01474 standard; protein; 835 AA.  
XX AC AAB01474;  
XX DT 20-OCT-2000 (first entry)  
XX DE Rat CARC (Cardiac related Ankyrin-Repeat Protein Kinase).  
XX Cardiac related ankyrin repeat protein kinase; CARC; cytoskeleton;  
XX cardiac cell growth factor receptor; cell differentiation; modulator;  
XX regulator; detection; cellular proliferation; cardiovascular disorder;  
XX heart failure; hypertension; cancer; sarcoma.  
XX OS Rattus norvegicus.  
XX PN WO200034330-A1.  
XX PD 15-JUN-2000.  
XX PF 10-DEC-1999; 99WO-US029465.  
XX PR 11-DEC-1998; 98US-0111938P.  
XX PR 14-APR-1999; 99US-00291839.  
XX PA (MILL-) MILLENNIUM PHARM INC.  
XX PI Raju J;  
XX WPI; 2000-431275/37.  
XX DR N-PSDB; AAA47608, AAA47609.  
XX New polynucleotide encoding cardiac-related ankyrin-repeat protein  
XX kinase, useful for treating disorders such as cardiovascular disorders,  
XX e.g. heart failure and cell differentiation disorders, e.g. cancer.  
XX Claim 1; Fig 5; 161pp; English.  
XX PS  
XX

CC CARK polypeptides are regulators of signal transmission from cellular  
CC receptors (e.g. cardiac cell growth factor receptor) and cytoskeleton  
CC function. They also act as modulators of cellular differentiation and  
CC cell death. Nucleic acids, or their fragments encoding CARK polypeptides  
CC are useful for detecting CARK nucleic acids especially mRNA, in a sample.  
CC CARK is useful for treating disorders associated with upregulation or  
CC downregulation of cellular proliferation such as, cardiovascular  
CC disorders (heart failure and hypertension) and disorders associated with  
CC cell differentiation such as cancer and sarcoma  
XX  
SQ

Sequence 835 AA;

Query Match 92.9%; Score 4079; DB 3; Length 835;  
Best Local Similarity 91.4%; Pred. No. 0;  
Matches 763; Conservative 38; Mismatches 34; Indels 0; Gaps 0;

QY 1 MGNYSRPTOTCTDEWKKVSESIVTIERLEDDLOIKKELTELNIIGSDEARSKVNL 60  
Db 1 MGNYSRPTOTCTDEWKKVSESIVTIERLEDDLOIKKELTELNIIGSDEARSKVNL 60  
QY 61 NYRTENGSLHLCCICGKKSHIRTLMLKGLRPSRLTRNGFTALHLAVYKDAELITSL 120  
Db 61 NYRTENGSLHLCCICGKKSHIRTLMLKGLRPSRLTRNGFTALHLAVYKDAELITSL 120  
QY 121 LHSAGDIOQVGGITLHIAHIAHIAHIAHIAHIAHIAHIAHIAHIAHIAHIAHIAH 180  
Db 121 LHSAGDIOQVGGITLHIAHIAHIAHIAHIAHIAHIAHIAHIAHIAHIAHIAHIAH 180  
QY 181 QVTRLLKFGADVNVSGRPLHLASAKGFLNIAKLMEEGSKADVNAQDNEDHVPVHL 240  
Db 181 QVTRLLKFGADVNVSGRPLHLASAKGFLNIAKLMEEGSKADVNAQDNEDHVPVHL 240  
QY 241 FCSRFGHDIVKYLQSDLEVPQHVNIYGTPTLHLACYNGKFEVAKIITQSGTSFLTK 300  
Db 241 FCSRFGHDIVKYLQSDLEVPQHVNIYGTPTLHLACYNGKFEVAKIITQSGTSFLTK 300  
QY 301 ENIFSETAFHSACTYKSIIDVAKFLLDQNVININHQGRDGTGLHSACVGHIRLVQFLL 360  
Db 301 ENIFSETAFHSACTYKSIIDVAKFLLDQNVININHQGRDGTGLHSACVGHIRLVQFLL 360  
QY 361 DNGADMLIVACDPSRSGKDEQTLWVAYEKGHDAIVTLLKHYKRPQDELPCEYSPQG 420  
Db 361 DNGADMLIVACDPSRSGKDEQTLWVAYEKGHDAIVTLLKHYKRPQDELPCEYSPQG 420  
QY 421 GDGSYVSVPSPLGKIKMTKEKADILLIRAGLPSHFLQISEIEFHEIIGSGSFGKVKYK 480  
Db 421 GDGSYVSVPSPLGKIKMTKEKADILLIRAGLPSHFLQISEIEFHEIIGSGSFGKVKYK 480  
QY 481 RCNRKIVAIKRYRANTYCSKSDVDMFCREVSILCOLNHPQCVIOFVGACLNDSQFAIVTQ 540  
Db 481 RCNRKIVAIKRYRANTYCSKSDVDMFCREVSILCOLNHPQCVIOFVGACLNDSQFAIVTQ 540  
QY 541 YISGSLPSLLHBOKRILDLQSLIIIVDVAKMEYHLNLTQPIIHRDLNSHNLILYEDG 600  
Db 541 YISGSLPSLLHBOKRILDLQSLIIIVDVAKMEYHLNLTQPIIHRDLNSHNLILYEDG 600  
QY 601 HAVVADFGESRFLQSDLEDNMTKQGNLRWMAPEVFTQCTRYTIKADVFSYALCLWEILT 660  
Db 601 HAVVADFGESRFLQSDLEDNMTKQGNLRWMAPEVFTQCTRYTIKADVFSYALCLWEILT 660  
QY 661 GEIPPAHLKPAADAAADWYHHRPIPIGYSTPKPISSLLIRGNWACPEGRPESEVVMKLE 720  
Db 661 GEIPPAHLKPAADAAADWYHHRPIPIGYSTPKPISSLLIRGNWACPEGRPESEVVMKLE 720  
QY 721 ECLCNIELMSPASSNSGSLSPSSSDCLVNRGPGRSHVAALRSFELEYALNARSYAA 780  
Db 721 ECLCNIELMSPASSNSGSLSPSSSDCLVNRGPGRSHVAALRSFELEYALNARSYAA 780  
QY 781 LSQAGQYSSOGLSLEEMKRSLOVTPIDKYGVSDPMSMHFSCNRSSEFDS 835  
Db 781 WSQSVGTHSNPGLSLEEMNRSTQYSTVDKYGVSDPMSLTHLSRQDQDNFEDSN 835

## RESULT 6

ABR41887

ID ABR41887 standard; protein; 835 AA.

XX ABR41887;  
AC ABR41887;

XX 11-AUG-2003 (first entry)

XX Rat cardiac-related ankyrin-repeat protein kinase (CARK).

XX Rat; CARK; cardiac-related ankyrin-repeat protein kinase; enzyme;  
KW cardiant; hypotensive; cytostatic.  
XX Rattus norvegicus.

XX Location/Qualifiers

FT Domain 66..99  
/note= "Ankyrin repeat domain"  
FT Domain 100..132  
/note= "Ankyrin repeat domain"  
FT Domain 133..165  
/note= "Ankyrin repeat domain"  
FT Domain 168..198  
/note= "Ankyrin repeat domain"  
FT Domain 199..233  
/note= "Ankyrin repeat domain"  
FT Domain 234..264  
/note= "Ankyrin repeat domain"  
FT Domain 269..302  
/note= "Ankyrin repeat domain"  
FT Domain 306..338  
/note= "Ankyrin repeat domain"  
FT Domain 339..371  
/note= "Ankyrin repeat domain"  
FT Region 411..415  
/note= "LXCXE motif"  
FT Domain 463..716  
/note= "protein kinase domain"

PN WO2003020912-A2.

XX 13-MAR-2003.

XX 04-SEP-2002; 2002WO-USO28300.

XX 05-SEP-2001; 2001US-00947199.

XX (MILL-) MILLENNIUM PHARM INC.

XX Raju J;

XX WPI: 2003-290188/28.  
XX N-PSDB: ACC48581.

XX Novel isolated cardiac-related ankyrin-repeat protein kinase polypeptide,  
XX useful for treating cellular growth related disorders which include  
XX cardiovascular disorders and proliferative and/or differentiative  
XX disorders.

XX Claim 14; Fig 5A-D; 158pp; English.

XX The present sequence is the protein sequence of a novel 96 kDa rat  
XX cardiac-related ankyrin-repeat protein kinase, designated CARK, which  
XX plays a role in signalling pathways associated with cardiac cell growth  
XX or differentiation. CARK is a dual specificity protein kinase having  
XX specificity for both serine/threonine and tyrosine residues. Human CARK  
XX (see ABR41886) expression is increased in ischaemic heart tissue samples,  
XX suggesting a role in the regulation of cardiac cell growth and/or  
XX differentiation and the pathogenesis of cardiovascular disorders, e.g.  
XX congestive heart failure and cardiac hypertrophy. The invention provides  
XX human and rat CARK proteins, fusion proteins, antigenic peptides and anti-  
XX CARK antibodies, and also CARK nucleic acids, antisense molecules,  
XX recombinant expression vectors, host cells and transgenic animals. These

are useful as modulating agents for regulating a variety of cellular processes, e.g. cardiac cellular process, for modulating the phosphorylation state of a CARP molecule or one or more proteins involved in cellular growth or differentiation, for modulating cell behaviour or as targets and therapeutic agents controlling cardiac cell proliferation, differentiation, hypertrophy and migration, for modulating intra- or inter-cellular signalling and/or gene transcription, for modulating cell proliferation, growth, differentiation, survival and/or migration, for regulating transmission of signals from cellular receptors, for modulating entry of cells, e.g. cardiac precursor cells, into mitosis, or for regulating cytoskeletal function. The proteins and nucleic acids are useful for treating cellular growth related disorders which include cardiovascular disorders (such as heart failure, hypertension), and proliferative and/or degenerative disorders (such as cancer). They are also useful in screening assays, detection assays (e.g. chromosomal mapping, tissue typing, forensic biology), predictive medicine (e.g., diagnostic assays, prognostic assays, monitoring clinical trials) and pharmacogenomics, and in methods of treatment. A CARP polypeptide can be used to generate a specific antibody, to screen for naturally-occurring CARP substrates, and to screen for drugs or compounds which modulate CARP activity. It is also useful as a bait protein in a yeast two-hybrid or three-hybrid assay and for identifying other proteins which bind to, or interact with, CARP or which are involved in CARP activity

Sequence 835 AA;

Query Match 92.9%; Score 4079; DB 6; Length 835;  
Best Local Similarity 91.4%; Pred. No. 0;  
Matches 763; Conservative 38; Mismatches 34; Indels 0; Gaps 0;

QY 1 MGNYSRPTCTDDEWKKVSVSYTIERLEDLQIKEKELTELRNIFGSDPAFKNVL 60  
DB 1 MGNYSRPTCTDDEWKKVSVSYTIERLEDLQIKEKELTELRNIFGSDPAFKNVL 60  
QY 61 NYTENGSLHLHCCTCGGKKSHIRFLMKLGRRLTRNGTALHLAVYKDNELTSL 120  
DB 61 NYRTERGLSLHLCCVCGGKSHIRFLMKLGRRLTRNGTALHLAVYKDNELTSL 120  
QY 121 LHSGADIQVGGGLTALHTATAGLEADVLQHGANNVQDAVFTPLHLAAYGHE 180  
DB 121 LHSGADVQVGGGLTALHTAAGPEAEVLLQHGANNVQDAVFTPLHLAAYGHE 180  
QY 181 QVTRLLKFGADVNSGEVDRPLHLASAKGFLNIAKLMEGSKADVNAQDNEDHVLPH 240  
DB 181 QVTSVLLKFGADVNSGEVDRPLHLASAKGFLNIAKLMEGSKADVNAQDNEDHVLPH 240  
QY 241 FCSRFEGHDIKVLQSDLEVPVHNYIYGDTPHLACVNGKFEVAKETIQISGESLTK 300  
DB 241 FCSRFEGHNIIVSLQSDLEVPVHNYIYGDTPHLACVNGKFEVAKETIQVGTGESLTK 300  
QY 301 ENIFSETAFHSACTYGSIDLKFLDQNVININHGQRDGHTGLHSACVHGHIRLVQELL 360  
DB 301 ENIFSETAFHSACTYGNIDLKFLDQNAVININHGQRDGHTGLHSACVHGHIRLVQELL 360  
QY 361 DNGADNMLVACDSRSRSGEKDEOTCLMAYEKHDAIVTLKHYKRPDELPFCNYSQPG 420  
DB 361 DNGADNMLVACDSRSRSGEKDEOTCLMAYEKHDAIVTLKHYKRPDELPFCNYSQPG 420  
QY 421 GDGSYVSPSPGLKISMTKEKADILLRAGLPSHFLQLSIEPHEIIGSGSGFKVYKG 480  
DB 421 GDGSYVSPSPGLKISMTKEKADVLLRAELPSRFLQLSIEPHEIIGSGSGFKVYKG 480  
QY 481 RCRNKIIVAKRYRANTYVCSKSDVDMFCREVSILCQLNHPCVITQVGCALNDFSPFAIVTQ 540  
DB 481 RCRNKIIVAKRYRANTYVCSKSDVDMFCREVSILCQLNHPCVITQVGCALNDFSPFAIVTQ 540  
QY 541 YISGGSLSLHLHQRKRLDLSKLITAVDVAKGMEVHLNLTQPIIHRDLNSHNLLEYDG 600  
DB 541 YISGGSLSLHLHQRKRLDLSKLITAVDVAKGMEVHLNLTQPIIHRDLNSHNLLEYDG 600  
QY 601 HAVVADPGEFRFLQSLDEDNMTKQGNLRWMAPEVFTQCTRTYIKADVPFVSLCWLWEIT 660  
DB 601 HAVVADPGEFRFLQSLDEDNMTKQGNLRWMAPEVFTQCTRTYIKADVPFVSLCWLWEIT 660

QY 661 GEIPFAHLKPAAAAADMAVYHRRPIGYSIKPISSLLIRGNACPEGRPFSEVVMKLE 720  
DB 661 GEIPFAHLKPAAAAADMAVYHRRPIGYSIKPISSLLIRGNACPEGRPFSEVVMKLE 720  
QY 721 ECLCNELMSPASSNSSLSPSSSDCLVNRGGPRSHVAALRSRFELEVALNARSAA 780  
DB 721 ECLCNELMSPASSNSSLSPSSSDCLVNRGGPRSHVAALRSRFELEVALNARSAG 780  
QY 781 LSGAGQYSSQGLSLEMKRSIQYTPIDKYGVSDPMSMHFSCNRSSEFSS 835  
DB 781 WSGVQTHSNPGLSLEEMNRSTQYSTDVKYGVSDPMSLTLHLSQDDSNFEDSN 835

RESULT 7  
AAE37962

ID AAE37962 standard; protein; 603 AA.

AC AAE37962;

XX 06-NOV-2003 (first entry)

XX Human kinase and phosphatase (KPP-7) protein.

XX Human; kinase; phosphatase; KPP; cell proliferative disorder; hepatitis;  
XX atherosclerosis; cirrhosis; haemoglobinuria; polycythaemia vera; cancer;  
XX psoriasis; thrombocytopenia; developmental disorder; Reiter's syndrome;  
XX renal tubular acidosis; anaemia; mental retardation; Alzheimer's disease;  
XX neurological disorder; Parkinson's disease; epilepsy; glomerulonephritis;  
XX autoimmune disorder; inflammatory disorder; AIDS; Goodpasture's disease;  
XX acquired immune deficiency syndrome; immunosuppressive; Crohn's disease;  
XX neotopic; transgenic; dermatitis; multiple sclerosis; diabetes mellitus;  
XX allergy; gout; Grave's disease; Hashimoto's thyroiditis; bowel syndrome;  
XX osteoporosis; rheumatoid arthritis; Sjogren's syndrome; ophthalmological;  
XX gene therapy; asthma; anticonvulsant; uropathic; pancreatitis.

XX Homo sapiens.

XX W02003050084-A2.

XX 19-JUN-2003.

XX 06-DEC-2002; 2002WO-US039126.

XX 07-DEC-2001; 2001US-0340235P.

XX 19-DEC-2001; 2001US-0343007P.

XX 21-DEC-2001; 2001US-0343546P.

XX 04-FEB-2002; 2002US-0354388P.

XX 15-FEB-2002; 2002US-0357675P.

XX (INCY-) INCYTE GENOMICS INC.

XX Kable AE, Chien D, Wilson AD, Swarnakar A, Gorvad AE;

XX Hafalia AJA, Emerling BM, Ramkumar J, Jin P, Griffin JA, Marquis JP;

XX Baughn MR, Chawla NK, Lehr-Mason PM, Khare R, Lee S, Hawkins PR;

XX Becha SD, Lee SY, Sprague WW, Zebbarjadian Y;

XX WPI; 2003-532894/50.

XX N-PSDB; AAD57334.

XX New human kinases and phosphatases and polynucleotides, useful for

XX diagnosing, treating or preventing autoimmune or inflammatory disorders

XX (e.g. AIDS, allergy or anemia), multiple sclerosis, osteoarthritis,

XX cancer or hepatitis.

XX Claim 1; Page 200-201, 282pp; English.

XX The invention relates to an isolated polypeptide, which is a human kinase

XX and phosphatase (KPP). KPP agonists and antagonists are useful for

XX diagnosing, treating or preventing disorders associated with aberrant

XX expression of KPP, particularly cell proliferative disorders (e.g.

XX arteriosclerosis, atherosclerosis, cirrhosis, hepatitis, paroxysmal

XX nocturnal haemoglobinuria, polycythaemia vera, psoriasis, primary

CC thrombocytopenia or cancer), developmental disorders (eg. renal tubular  
CC acidosis, anaemia or mental retardation), neurological disorders (e.g.  
CC Alzheimer's disease, Parkinson's disease or epilepsy), autoimmune/  
CC inflammatory disorders (e.g. AIDS; acquired immune deficiency syndrome,  
CC allergies, asthma, autoimmune thyroiditis, contact dermatitis, Crohn's  
CC disease, diabetes mellitus, glomerulonephritis, Goodpasture's syndrome,  
CC gout, Grave's disease, Hashimoto's thyroiditis, irritable bowel syndrome,  
CC multiple sclerosis, osteoarthritis, osteoporosis, pancreatitis, Reiter's  
CC syndrome, rheumatoid arthritis, Sjogren's syndrome, uveitis), or viral,  
CC bacterial, fungal, parasitic, protozoan or helminthic infections. The KPP  
CC is useful in assessing the effects of exogenous compounds on the  
CC expression of nucleic acids and kinases and phosphatases. KPP gene is  
CC useful in gene therapy and for creating transgenic animals to model human  
CC disease. The present sequence is human KPP protein  
XX  
SQ Sequence 603 AA;

Query Match 70.7%; Score 3102; DB 6; Length 603;  
Best Local Similarity 100.0%; Pred. No. 1e-276;  
Matches 590; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MGNYSRPTQCTDEWKKKSESYYTIERLEDLQIKKELTELRLNIFGSDFAFSKVL 60  
DB 1 MGNYSRPTQCTDEWKKKSESYYTIERLEDLQIKKELTELRLNIFGSDFAFSKVL 60  
QY 61 NYRTENGSLHLACICGGKSHRTLMKGLRSLTRNGFTALHLAVYKDAEILTSI 120  
DB 61 NYRTENGSLHLACICGGKSHRTLMKGLRSLTRNGFTALHLAVYKDAEILTSI 120  
QY 121 LHSGLADIQOQVGGITLHATIAHLEAADVLLQHGANNVNIQDAVFPTPLHIAAYGHE 180  
DB 121 LHSGLADIQOQVGGITLHATIAHLEAADVLLQHGANNVNIQDAVFPTPLHIAAYGHE 180  
QY 181 QVTRLLKFGADVNSGEVGRPLHLASAKGFLNIAKLMEBSKADVNAQDNEDHVP 240  
DB 181 QVTRLLKFGADVNSGEVGRPLHLASAKGFLNIAKLMEBSKADVNAQDNEDHVP 240  
QY 241 FCSRFGRHDIYKYLQSDLEVPQHVNIYGDTPHLACYNKFEVAKIITQISGTS 300  
DB 241 FCSRFGRHDIYKYLQSDLEVPQHVNIYGDTPHLACYNKFEVAKIITQISGTS 300  
QY 301 ENIFSETAFHSACTYKGSIDLKFLDQNVININHQGRDGHGTHLSACYHGIRLV 360  
DB 301 ENIFSETAFHSACTYKGSIDLKFLDQNVININHQGRDGHGTHLSACYHGIRLV 360  
QY 361 DNGADNVUACDPSRSSEKDEQCLMWAYEKHGDAIVTLKHYKRPQDELPCNEYS 420  
DB 361 DNGADNVUACDPSRSSEKDEQCLMWAYEKHGDAIVTLKHYKRPQDELPCNEYS 420  
QY 421 GDGSYVSPPLGKIKMTKEKADILLRAGLPSHFHLQSEIEFHEIIGSGSFGK 480  
DB 421 GDGSYVSPPLGKIKMTKEKADILLRAGLPSHFHLQSEIEFHEIIGSGSFGK 480  
QY 481 RCNRKIVAKRYRANTYCSKSDVDMFCREVSILCOLNHCPCVIOFGACLNDS 540  
DB 481 RCNRKIVAKRYRANTYCSKSDVDMFCREVSILCOLNHCPCVIOFGACLNDS 540  
QY 541 YISGSLPSLLHCKRILDLQSLKIIIVDVAKMGWEYLHNLITQPIIHRDLN 590  
DB 541 YISGSLPSLLHCKRILDLQSLKIIIVDVAKMGWEYLHNLITQPIIHRDLN 590

RESULT 8  
AAU20525  
ID AAU20525 standard; protein; 148 AA.  
XX  
AC AAU20525;  
XX  
DT 06-DEC-2001 (first entry)  
XX  
DE Human secreted protein, Seq ID No 517.  
XX  
KW Immunomodulatory; human immunodeficiency virus; HIV; anaemia; angina;

KW rheumatoid arthritis; antiarteriosclerotic; cardiac; vascular;  
KW cerebroprotective; thrombolytic; antimicrobial; ophthalmological;  
KW cystic; Alzheimer's disease; Parkinson's disease; human; cancer;  
KW multiple sclerosis; cancer; hyperproliferative disorder; infection;  
KW Gaucher's disease; neurological disease; cerebrovascular disorder;  
KW thrombosis; wound healing.  
XX  
OS Homo sapiens.  
XX  
FN WO200155326-A2.  
XX  
PD 02-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US001347.  
XX  
PR 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225575P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226688P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.

XX	Immunomodulatory; human immunodeficiency virus; HIV; anaemia; angina;
KW	rheumatoid arthritis; antiarteriosclerotic; cardiant; vascular;
KW	cerbroprotective; thrombolytic; antimicrobial; ophthalmological;
KW	cystostatic; Alzheimer's disease; Parkinson's disease; human; cancer;
XX	

KW multiple sclerosis; cancer; hyperproliferative disorder; infection;  
KW Gaucher's disease; neurological disease; cerebrovascular disorder;  
KW thrombosis; wound healing.

OS Homo sapiens.

XX WO200155326-A2.

PN

PD

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 26-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.



```
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144086P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 21-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145234P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 28-JUL-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 06-AUG-1999; 99US-0147419P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 16-AUG-1999; 99US-0149375P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149823P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.

PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159283P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159337P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161933P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 11.2%; Score 493; DB 3; Length 589;
Best Local Similarity 31.6%; Pred. No. 7.6e-36;
Matches 136; Conservative 86; Mismatches 145; Indels 64; Gaps 15;

QY 416 YSOPGGDSYVSPSLGKIKSM--TKKA-----DILLRAGLPS----HFHL 458
DB 51 WSQETDGLRDALSKILKLDQPGSKQKSIFFEHDKSSNELIPACIPIPTDGTDEWEI 110
QY 459 QLSIEFHEITIGSGFGKVKYGCNRKNIKVAIKRYRANTYCSKSDVDMPCREVSILCOLNH 518
DB 111 DVTQKIEKKVAGSYGDLHRTYCSQVIAIKLPDR--VNNEMLEFESQEVFIRKVRH 169
QY 519 PCVIOFVACINDPSQFAIVTQYISGSLFSLHCKRILDIQSKLITAVDVAKMEYLH 578
DB 170 KNVQFLGACTRSPT-LCIVTFEMARGSIYDFLHKQKCAFQLOTLKVALDVAKMSYLH 228
QY 579 NLTOPIIHRDLNSHNILYEDGHAVVADFGEGRFLOSLEDNNTKQPNLRWMAPEVTQ 638
DB 229 Q--NNIHRDLKTNLLMDEHGLVKVADFGVARV--QIESGVMTAETGYRWMAPEVI-E 283
QY 639 CTRYTIKADVPFALCLWEILTCETIPFAHLKPAADAMAYHHIRPPIGYSIPKPISSIL 698
DB 284 HKPYNHKADVFVAIVLWELLTGDIPYAFITPQAAVGVQVGLRKPDKPKTKHVKVGLL 343
QY 699 IRGNACPEGRPEFSEVMKLEBCLNIELMSPASSNSGSLG-----PSSSD 747
DB 344 ERCWHQDPEORPLFEETIEMLQOI-----MKEPVTVFGSATAVEEMVFLSWGRFSSEQ 398
QY 748 CLV-----NRGGPGRSHVAALRSRFELE---YALNARSVAALSQAGQYSSQG 792
DB 399 QQVINKTGTFFNYDNKYRGVSSRS-IAKLKEDSEIDKXGFLIN---HARVLVSGRE---- 450
QY 793 LSLEMKRSIQ 803
```

```
Db      451  -SYEKGKALQ 460
RESULT 11
AAG45983
ID  AAG45983 standard; protein; 732 AA.
XX
AC  AAG45983;
XX
DT  18-OCT-2000 (first entry)
XX
DE  Arabidopsis thaliana protein fragment SEQ ID NO: 57799.
XX
KW  Protein identification; signal transduction pathway; metabolic pathway;
KW  hybridisation assay; genetic mapping; gene expression control; promoter;
XX  termination sequence.
OS  Arabidopsis thaliana.
PN  EP1033405-A2.
FD  06-SEP-2000.
XX
XX  25-FEB-2000; 2000EP-00301439.
XX
XX  25-FEB-1999; 99US-0121825P.
PR  18-JUN-1999; 99US-0139455P.
PR  18-JUN-1999; 99US-0139456P.
PR  18-JUN-1999; 99US-0139457P.
PR  18-JUN-1999; 99US-0139458P.
PR  18-JUN-1999; 99US-0139459P.
PR  18-JUN-1999; 99US-0139460P.
PR  18-JUN-1999; 99US-0139461P.
PR  18-JUN-1999; 99US-0139462P.
PR  18-JUN-1999; 99US-0139463P.
PR  18-JUN-1999; 99US-0139750P.
PR  18-JUN-1999; 99US-0139763P.
PR  21-JUN-1999; 99US-0139817P.
PR  22-JUN-1999; 99US-0139899P.
PR  23-JUN-1999; 99US-0140353P.
PR  23-JUN-1999; 99US-0140354P.
PR  24-JUN-1999; 99US-0140695P.
PR  28-JUN-1999; 99US-0140823P.
PR  29-JUN-1999; 99US-0140991P.
PR  30-JUN-1999; 99US-0141287P.
PR  01-JUL-1999; 99US-0141842P.
PR  01-JUL-1999; 99US-0142154P.
PR  02-JUL-1999; 99US-0142055P.
PR  06-JUL-1999; 99US-0142390P.
PR  08-JUL-1999; 99US-0142803P.
PR  09-JUL-1999; 99US-0142920P.
PR  12-JUL-1999; 99US-0142977P.
PR  13-JUL-1999; 99US-0143542P.
PR  14-JUL-1999; 99US-0143624P.
PR  15-JUL-1999; 99US-0144005P.
PR  16-JUL-1999; 99US-0144085P.
PR  16-JUL-1999; 99US-0144086P.
PR  19-JUL-1999; 99US-0144325P.
PR  19-JUL-1999; 99US-0144331P.
PR  19-JUL-1999; 99US-0144332P.
PR  19-JUL-1999; 99US-0144333P.
PR  19-JUL-1999; 99US-0144334P.
PR  19-JUL-1999; 99US-0144335P.
PR  20-JUL-1999; 99US-0144352P.
PR  20-JUL-1999; 99US-0144632P.
PR  20-JUL-1999; 99US-0144884P.
PR  21-JUL-1999; 99US-0144814P.
PR  21-JUL-1999; 99US-0145086P.
PR  21-JUL-1999; 99US-0145088P.
PR  22-JUL-1999; 99US-0145085P.
PR  22-JUL-1999; 99US-0145087P.
PR  22-JUL-1999; 99US-0145089P.
PR  22-JUL-1999; 99US-0145192P.
PR  23-JUL-1999; 99US-0145145P.
PR  23-JUL-1999; 99US-0145218P.
PR  23-JUL-1999; 99US-0145224P.
PR  26-JUL-1999; 99US-0145276P.
PR  27-JUL-1999; 99US-0145913P.
PR  27-JUL-1999; 99US-0145918P.
PR  27-JUL-1999; 99US-0145919P.
PR  28-JUL-1999; 99US-0145951P.
PR  02-AUG-1999; 99US-0146386P.
PR  02-AUG-1999; 99US-0146388P.
PR  02-AUG-1999; 99US-0146389P.
PR  03-AUG-1999; 99US-0147038P.
PR  04-AUG-1999; 99US-0147204P.
PR  04-AUG-1999; 99US-0147302P.
PR  05-AUG-1999; 99US-0147303P.
PR  05-AUG-1999; 99US-0147260P.
PR  06-AUG-1999; 99US-0147303P.
PR  06-AUG-1999; 99US-0147416P.
PR  09-AUG-1999; 99US-0147493P.
PR  09-AUG-1999; 99US-0147935P.
PR  10-AUG-1999; 99US-0148171P.
PR  11-AUG-1999; 99US-0148319P.
PR  12-AUG-1999; 99US-0148341P.
PR  13-AUG-1999; 99US-0148565P.
PR  13-AUG-1999; 99US-0148684P.
PR  16-AUG-1999; 99US-0149368P.
```

```
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149732P.
PR 20-AUG-1999; 99US-0149733P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151085P.
PR 27-AUG-1999; 99US-0151086P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154039P.
PR 16-SEP-1999; 99US-0154779P.
PR 20-SEP-1999; 99US-0155139P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158023P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161408P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 11.2%; Score 493; DB 3; Length 732;
Best Local Similarity 31.6%; Pred. No. 1.1e-35;
Matches 136; Conservative 86; Mismatches 145; Indels 64; Gaps 15;

QY 416 YSQPGDGSYVSPFLGKISM--TKEA-----DILLRAGLPS---HFHL 458
DB 194 WQSEETDGLRDLSKEILKDKDQPGSKQKISFPEHDKSNSELIPACIEIPTDGTDEWEI 253
QY 459 QLSSEIEFHEIIIGSGFGKYKGRCKNIKIVAKRYRANTYCSKSDVMFCREVSIQLCLNH 518
DB 254 DVTQLKIEKVASGSGDLHRTYCSQEVAIKFLKPDPR-VNNEMLREFSQEVFTMRKVRH 312
```

```
QY 519 PCVIOFVGACLNDPQFAIVTQVYISGSLFSLHROKRILOSLKLIITAVDAKMEYLH 578
DB 313 KNVVQFLGACTRSPT-LCIVTFMARGSYDFLHKQKCAFKLQTLKVALDVAKGMSYLH 371
QY 579 NLTPQIHRDLNSHNILLYEDGHAVVADFGESRFLQSLDEDNWTQPGNLRWMAPEVTO 638
DB 372 Q--NNIHRDLKTANLMDHGLVKVADFGVARV--QIESGVMTAETGYRWMAPEVI-E 426
QY 639 CTRYTIKADYFVYALCLWEILTGEIPFAHLKPAATAADWAYHHIRPPYISPKPISILL 698
DB 427 HKPYNHKADYFVYALVWELLTGDIPYAFLLTPLQAAVGVVQGLRKPPIKTHPKVKGLL 486
QY 699 IRGNACPEGRPEFSEVWKLBECLCNIELMSPASSNSSGSL-----PSSSSD 747
DB 487 ERCWHODPEORPLFEBIEMLOOI-----MKEPVTVFGSIAVEMVFLSWGRFSSBQQ 541
QY 748 CLV-----NRGGPGRSHVAALRSRFELE---YALNARSYAALSQAGYSSQG 792
DB 542 QQVINKTGTENYDNKYRGVSSRS-IAKIKEDSEIDKDGFLN---HARVLVSGRE---- 593
QY 793 LSUEEMKRSIQ 803
DB 594 -SVEKGGKALQ 603

RESULT 12
AAG45982
ID AAG45982 standard; protein; 760 AA.
XX AAG45982;
XX 18-OCT-2000 (first entry)
XX Arabidopsis thaliana protein fragment SEQ ID NO: 57798.
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
OS Arabidopsis thaliana.
XX EPI033405-A2.
XX 06-SEP-2000.
XX 25-FEB-2000; 2000EP-00301439.
XX 25-FEB-1999; 99US-0121825P.
XX 05-MAR-1999; 99US-0123180P.
XX 09-MAR-1999; 99US-0123548P.
XX 23-MAR-1999; 99US-0125788P.
XX 25-MAR-1999; 99US-0126264P.
XX 29-MAR-1999; 99US-0126785P.
XX 01-APR-1999; 99US-0127452P.
XX 06-APR-1999; 99US-0128234P.
XX 08-APR-1999; 99US-0128714P.
XX 16-APR-1999; 99US-0129845P.
XX 19-APR-1999; 99US-0130077P.
XX 21-APR-1999; 99US-0130449P.
XX 23-APR-1999; 99US-0130510P.
XX 23-APR-1999; 99US-0130891P.
XX 28-APR-1999; 99US-0131449P.
XX 30-APR-1999; 99US-0132048P.
XX 30-APR-1999; 99US-0132407P.
XX 04-MAY-1999; 99US-0132484P.
XX 05-MAY-1999; 99US-0132485P.
XX 06-MAY-1999; 99US-0132486P.
XX 06-MAY-1999; 99US-0132487P.
XX 11-MAY-1999; 99US-0132863P.
XX 14-MAY-1999; 99US-0134218P.
XX 14-MAY-1999; 99US-0134219P.
XX 14-MAY-1999; 99US-0134221P.
```



PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161922P.  
PR 28-OCT-1999; 99US-0161933P.  
PR 29-OCT-1999; 99US-0162142P.  
Query Match 11.2%; Score 493; DB 3; Length 760;  
Best Local Similarity 31.6%; Pred. No. 1.2e-35;  
Matches 136; Conservative 86; Mismatches 145; Indels 64; Gaps 15;  
QY 416 YSQPGCGSGYVSFPLGKISM--TKEKA-----DILLRAGLPS----HFHL 458  
DB 222 WQETDGLRALSKELKLDQPGSKQKSIFFHDKSSNELIPACIEIPTDGTDEWEI 281  
QY 459 QLSIEFHEIIGSGFQKVKYGRCKNIVAKRYRANTYCSKSDVDMFCREVSILCOLNH 518  
DB 282 DVTQLKIEKVASGSGDLHRTYCSQEVAKFLKPD--VNNEMLEFSGQEVFIMKVRH 340  
QY 519 PCVIOFVACLNDSPFAIVQYISGSLFSLHEQKRILDSKLIIVADVAKGMEYLH 578  
DB 341 KVVQFLGACTRSTP-LCIVTFWARGSIYDFLHKQKCAFLOTLLKVALDVAKGMYLH 399  
QY 579 NLQPTIHRDLSNHLIYEDGHAVADFGESRFLQSLDDEMTKQGNLRWMAPEVFTQ 538  
DB 400 Q--NNIIHRDLKTNALLMDEHGLVKVADFGVARY--QIESGVMTAETGTYYRWMAPEVI-E 454  
QY 639 CTRYTKADVSFALCWEIITGTETPAHLKPAAADMAVHHIRPPDIGYSIKPISL 698  
DB 455 HKPYNHKAQVSFALVWELLTGDIPTAFLTPLOAVGVQKGLRKPDKPTKPKVKGLL 514  
QY 699 IRGNWACPEGRPEFSEVVMKLECLNIELMSPASSNSGSL-----PSSSD 747  
DB 515 ERCWQDPEQPLPEEIIEMLQI-----MKEPTVFGSASIAVEEMVFLSGRPSSEQ 569  
QY 748 CLV-----NRGPGRSVHAALRSRFELE---YALNARSVAALQSAGQYSSQG 792  
DB 570 QQVINKTGTFFNYDKYRGVSSRS-TAKLKEDSEIDKDGFLIN---HARVLVSGRE---- 621  
QY 793 LSLSEMKSLQ 803  
DB 622 -SYEKGKALQ 631

RESULT 13  
ADD47763  
ID ADD47763 standard; protein; 1881 AA.  
XX AC ADD47763;  
XX DT 29-JAN-2004 (first entry)  
XX DE Human Protein XP\_016774, SEQ ID NO 13459.  
XX KW Human; pain; neuronal tissue; gene therapy;  
XX KW spinal segmental nerve injury; chronic constriction injury; CCI;  
XX KW spared nerve injury; SNI; Chung.  
XX OS Homo sapiens.  
XX PN WO2003016475-A2.  
XX PD 27-FEB-2003.  
XX PF 14-AUG-2002; 2002WO-US025765.  
XX PR 14-AUG-2001; 2001US-0312147P.  
XX PR 01-NOV-2001; 2001US-0346382P.  
XX PR 26-NOV-2001; 2001US-0333347P.  
XX PA (GENO ) GEN HOSPITAL CORP.

PA (FARB ) BAYER AG.  
XX PI Woolf C, D'urso D, Befort K, Costigan M;  
XX WPI; 2003-268312/26.  
DR GENBANK; XP\_016774.  
XX New composition comprising two or more isolated polypeptides, useful for  
PT preparing a medicament for treating pain in an animal.  
XX Claim 1; Page; 1017pp; English.  
XX The invention discloses a composition comprising two or more isolated rat  
CC or human polynucleotides or a polynucleotide which represents a fragment,  
CC derivative or allelic variation of the nucleic acid sequence. Also  
CC claimed are a vector comprising the novel polynucleotide, a host cell  
CC comprising the vector, a method for identifying a nucleotide sequence  
CC which is differentially regulated in an animal subjected to pain and a  
CC kit to perform the method, an array, a method for identifying an agent  
CC that increases or decreases the expression of the polynucleotide sequence  
CC that is differentially expressed in neuronal tissue of a first animal  
CC subjected to pain, a method for identifying a compound which regulates  
CC the expression of a polynucleotide sequence which is differentially  
CC expressed in an animal subjected to pain, a method for identifying a  
CC compound that regulates the activity of one or more of the  
CC polynucleotides, a method for producing a pharmaceutical composition, a  
CC method for identifying a compound or small molecule that regulates the  
CC activity in an animal of one or more of the polypeptides given in the  
CC specification, a method for identifying a compound useful in treating  
CC pain and a pharmaceutical composition comprising the one or more  
CC polypeptides or their antibodies. The polynucleotide or the compound that  
CC modulates its activity is useful for preparing a medicament for treating  
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction  
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
CC therapy). The sequence presented is a human protein (shown in Table 2 of  
CC the specification) which is differentially expressed during pain. Note:  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic form directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.  
XX Sequence 1881 AA;  
Query Match 10.5%; Score 460.5; DB 7; Length 1881;  
Best Local Similarity 27.6%; Pred. No. 5.2e-32;  
Matches 166; Conservative 76; Mismatches 176; Indels 183; Gaps 23;  
QY 25 VITIERLED---DLQIEK-ELTEL-----RNIFGSDEAFSKYNLV-----RTENGLSL 70  
DB 252 VIMVRLLDGRGAQIETKTDELTPHCAARN---GHVRISEILLDHGAPIQAKTKNGLS 308  
QY 71 LHLG-----CI-----CGGKSHIRTMILKGLRPSRL 97  
DB 309 IHMAAQGDHLDVRLQLQYDAEIDITLDHLTPHVAACHGHRV-AKVLDDGAKPNSR 367  
QY 98 TRNGFTALHAYKDNALITSLHSCADIQQVGYGLTALTATATAGHLEADVLQHG 157  
DB 368 ALNGFTPLHIAKCKHVRVNMELLLKTGASIDAVTESGLTPHVASFMGHLPIVKMLLQ 427  
QY 158 ANVNIQDAVFPTPLHIAAYYGHEQVTRLLKFGADVNVSGEVGDRFLHLASAKGFLNTAK 217  
DB 428 ASPNVSNKVTETPLHMAARAGHTEVAKYLLQNKAKVNAKAKDDQDTLHCAARIGHTNMV 487  
QY 218 LMEEGSKADVNAQDNEDHVPHFCSRFGH-----HDIVK 252  
DB 488 LLLLE--NNANPNLATTAGHTPLHIAAREGHVETVLALLEKEASQACMCKKGTPLHLVA 545  
QY 253 Y-----LLOSD-----LEV-----OPHYVNIYGDTP 273  
DB 546 YGKRVVAELLERDAHPNAGKNGLTPHVAHVHNNLDIVKLLPRGSGSPHSPANNGYTP 605  
QY 274 LHLACYNGKFEVAKIITQISGTSLSLTKENIFSTAFHSACTYKGSIDLKVLDDQNV--- 330  
DB 606 LHIAAKQNVQEVARSLLQYGG--SANAESVQGVTPFLHAAQEGHA-EMVALLLSKOANGN 662

QY 331 -----ININHQGRDGHGTHGACVYGHIRLVQFLD 361  
 Db 663 LGNKGSLTPLHLVAQEGHVPADVLIKHGMVMDATTRMGYTPLHVASHYGNIKLVKFLQ 722  
 QY 362 NGADMNLVACDPSRSSEKDEQCLMWAYEKGHDAIVT--LLKHVKRPQDELFCNYSOPG 420  
 Db 723 HQADV-----AKTKLGYSPLHQAAQOQHTDIVTLLKNGASP-----NEVSDG 767  
 QY 421 GDG-----SYVSPSPGLKIKSMTEKADILL--RAGLP SHF--LQSETEFHEI 468  
 Db 768 TTPLAIKRLGISVTDVL---KVVTDTSFVLVSDKHRMSPPTVDEILDVSEDEGEL 824  
 QY 469 I 469  
 Db 825 I 825

RESULT 14  
 ABU11523  
 ID ABU11523 standard; protein; 1536 AA.  
 AC ABU11523;  
 XX  
 DT 12-FEB-2003 (first entry)  
 XX  
 DE Human MDDT polypeptide SEQ ID 470.  
 XX  
 KW MDDT; human; disease detection and treatment molecule polypeptide;  
 KW anti-inflammatory; immunosuppressive; osteopathic; cytostatic; anti-HIV;  
 KW haemostatic; nephrotropic; antianaemic; antipsoriatic; hepatotropic;  
 KW gene therapy; protein replacement therapy; cell proliferative disorder;  
 KW cancer; adenocarcinoma; leukaemia; lymphoma; melanoma; sarcoma;  
 KW anaemia; Crohn's disease; acquired immunodeficiency syndrome; AIDS;  
 KW Goodpasture's syndrome; inflammation; osteoporosis; thrombocytopaenia;  
 KW psoriasis; hepatitis.  
 XX  
 OS Homo sapiens.  
 XX  
 FN WO200279449-A2.  
 XX  
 PD 10-OCT-2002.  
 XX  
 PF 27-MAR-2002; 2002WO-US009944.  
 XX  
 PR 28-MAR-2001; 2001US-0279619P.  
 PR 29-MAR-2001; 2001US-0280067P.  
 PR 29-MAR-2001; 2001US-0280068P.  
 PR 16-MAY-2001; 2001US-0291280P.  
 PR 17-MAY-2001; 2001US-0291829P.  
 PR 17-MAY-2001; 2001US-0291849P.  
 PR 19-JUN-2001; 2001US-0299428P.  
 PR 20-JUN-2001; 2001US-0299776P.  
 PR 20-JUN-2001; 2001US-0300001P.  
 XX  
 PA (INCYTE GENOMICS INC.  
 XX  
 PI Dafo A, Jones AL, Tran AB, Dahl CR, Gietzen D, Chinn J;  
 PI Dufour GS, Hillman JL, Yu JY, Tuason O, Yap PE, Amshey SR;  
 PI Daugherty SC, Dam TC, Liu TF, Nguyen DA, Kleeefeld Y, Gerstin EH;  
 PI Peralta CH, David MH, Lewis SA, Chen AJ, Panzer SR, Harris B;  
 PI Flores V, Marwaha R, Lo A, Lan RY, Urashka ME;  
 XX  
 DR WPI; 2003-058431/05.  
 DR N-PSDB; ABX34513.  
 XX  
 PT New purified disease detection and treatment molecule proteins and  
 PT polynucleotides, useful for diagnosing, treating or preventing cancers  
 PT (e.g. leukemia or sarcoma), anemia, Crohn's disease, AIDS, osteoporosis  
 PT or hepatitis.  
 XX  
 XX Claim 27; SEQ ID NO 470; 339pp + Sequence Listing; English.

CC This invention describes a novel disease detection and treatment molecule  
 CC polypeptide (MDDT) which has anti-inflammatory, immunosuppressive,  
 CC osteopathic, cytostatic, anti-HIV, haemostatic, nephrotropic,  
 CC antianaemic, antipsoriatic and hepatotropic activity. The polynucleotides  
 CC and the polypeptides of the invention can be used for gene therapy,  
 CC protein replacement therapy and are useful for treating a variety of  
 CC diseases or conditions. These polypeptides or polynucleotides are  
 CC particularly useful for diagnosing, treating or preventing cell  
 CC proliferative disorders (e.g. cancers including adenocarcinoma,  
 CC leukaemia, lymphoma, melanoma, myeloma or sarcoma), anaemia, Crohn's  
 CC disease, acquired immunodeficiency syndrome (AIDS), Goodpasture's  
 CC syndromes, inflammation, osteoporosis, thrombocytopaenia, psoriasis or  
 CC hepatitis. ABU11450-ABU11845 represent the MDDT polynucleotides encoded  
 CC by ABU11450-ABU11845, described in the disclosure of the invention. NOTE:  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 XX Sequence 1536 AA;

Query Match 10.3%; Score 454; DB 6; Length 1536;  
 Best Local Similarity 28.3%; Pred. No. 1.5e-31;  
 Matches 151; Conservative 70; Mismatches 175; Indels 138; Gaps 17;  
 QY 45 LRNIFGSDEAFSKVNLNRYTENGSLLLHLCICCGKKSHIRTLMLKGLRPSRLTRNGFTA 104  
 Db 48 VRLLOYDAEIDDIIDH-----LTPHVAHCGHHRV-AKVLLOKGAKNRSLALNGFT 101  
 QY 105 LHLAVYKDNALITSLHSGADIQQVGGGLTALHIAIAGHLEAADVLLQGANVITQD 164  
 Db 102 LHTACKNHRVNMELLTKGASIDAVTESGLTFLHVAFMGHLPTVKLLQORGASPNYSN 161  
 QY 165 AVFTFTLHIAAYVGHBOVTRLLKFCADVNSGEGVDRPLHLASAKGFINTAKLMEEGS 224  
 Db 162 VKVETPLHMAARAGHTEVAKYLLQNKAKVNAKDDQTPHCAARIGHTNMVKLLLE--N 219  
 QY 225 KADVNAQDNEDHVPHFCSRFH-----HDIVKY----- 253  
 Db 220 NNPNLATTAGHTPLHIAAREGHVETVLALLEKEASQACMTKKGTPLHVAAYKGRVA 279  
 QY 254 --LLOSD-----LEV-----QPHVVNIYGDTPHLHACYN 280  
 Db 280 ELLERDAHPNAGKNGLTPLHVAVHHNLDIVKLLFRGSGSPHSPAWNVTPLHIAKQ 339  
 QY 281 GKFEVAKIIOISGTESLTENIFSETAFHSACTYKSIDLVKFLLDQNV----- 330  
 Db 340 NOVEVAHSLLOYG--SANAESVQGVTPHLAAQEGHA-EMVALLSKQANGNLKNSGL 396  
 QY 331 -----ININHQGRDGHGTHGACVYGHIRLVQFLDNGADMNL 368  
 Db 397 TPLHLVAQEGHVPADVLIKHGMVMDATTRMGYTPLHVASHYGNIKLVKFLQHQADYN- 455  
 QY 369 VACDPSRSSEKDEQCLMWAYEKGHDAIVT--LLKHVKRPQDELFCNYSOPGSDG--- 423  
 Db 456 -----AKTKLGYSPLHQAAQOQHTDIVTLLKNGASP-----NEVSDGTTPLAIA 501  
 QY 424 ---SYVSPSPGLKIKSMTEKADILL--RAGLP SHF--LQSETEFHEI 469  
 Db 502 KRLGISVTDVL---KVVTDTSFVLVSDKHRMSPPTVDEILDVSEDEGELI 552

RESULT 15  
 ABB67412  
 ID ABB67412 standard; protein; 1549 AA.  
 XX  
 AC ABB67412;  
 XX  
 DT 26-MAR-2002 (first entry)  
 XX  
 DE Drosophila melanogaster polypeptide SEQ ID NO 29028.  
 XX  
 KW Drosophila; developmental biology; cell signalling; insecticide;  
 KW pharmaceutical.

XX	Drosophila melanogaster.
OS	
XX	WO200171042-A2.
FN	
XX	
XX	27-SEP-2001.
PD	
XX	
XX	23-MAR-2001; 2001WO-US009231.
PF	
XX	
XX	23-MAR-2000; 2000US-0191637P.
PR	
XX	11-JUL-2000; 2000US-00614150.
XX	
XX	(PEKE ) PE CORP NY.
PA	
XX	
XX	Venter JC, Adams M, Li PWD, Myers EW;
PI	
XX	
XX	WPI; 2001-656680/75.
DR	
DR	N-PSDB; ABU11515.
XX	
XX	New isolated nucleic acid detection reagent for detecting 1000 or more
PT	genes from Drosophila and for elucidating cell signaling and cell-cell
PT	interactions.
XX	
XX	Disclosure; SEQ ID NO 23028; 21pp + Sequence Listing; English.
PS	
XX	
XX	The invention relates to an isolated nucleic acid detection reagent
CC	capable of detecting 1000 or more genes from Drosophila. The invention is
CC	useful in developmental biology and in elucidating cell signalling and
CC	cell-cell interactions in higher eukaryotes for the development of
CC	insecticides, therapeutics and pharmaceutical drugs. The invention
CC	discloses genomic DNA sequences (ABU16176-ABU30511), expressed DNA
CC	sequences (ABU01840-ABU16175) and the encoded proteins (ABB55773-
CC	ABB572072). The sequence data for this patent did not form part of the
CC	printed specification, but was obtained in electronic format directly
CC	from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX	
XX	Sequence 1549 AA;
XX	

442	Qy	KADILLRAG----	LPSPFHQLQSLSEIPEHIIIGSGFGKVKYKGRCKNIWA-IKRYR	493
839	Db	LMQETLLSDSDSDSCDDLLDHNH-----	YKNATDDELKANTYQQQ	878
494	Qy	ANTYCKSDVDMFCREVSILCOLNHPCVIQFVGACL-----	NDPSOFAIVTQYISGGSL	547
879	Db	KNPDTTNTDHDL--TDVSV-----	LNNKELLPNWMSCIELTEIGHKPDNVVIARSOVHLGFL	933
548	Qy	FSLLEHKEKRLDLQSLKLIIVDVAKG--MEYLHNLTQPII-----	HR	587
934	Db	VSEFL-----	VDAGGSMRGYRHNGVLIIVPKCAEPTRICTRYVKPQR	977
588	Qy	DLSNHNILYEDGHAVADFGESFLQSLDDEDDNNTKQP-----	GNLR	629
978	Db	VNNPPPLM--EGEALV-----	SRAILEMSPDGMFLSPITLVEPHYGTLL	1019

Search completed: September 2, 2004, 14:21:15  
Job time : 151 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 2, 2004, 14:18:48 ; Search time 34 seconds  
(without alignments)  
1267.873 Million cell updates/sec

Title: US-10-626-173-2

Perfect score: 4390

Sequence: 1 MGNYKSRPTQCTDEWKKV.....PMSSMHFHSRNSSEFSS 835

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:\*

- 1: /cgn2\_6/prodata/2/iaa/5A COMB.pep:\*
- 2: /cgn2\_6/prodata/2/iaa/5B COMB.pep:\*
- 3: /cgn2\_6/prodata/2/iaa/6A COMB.pep:\*
- 4: /cgn2\_6/prodata/2/iaa/6B COMB.pep:\*
- 5: /cgn2\_6/prodata/2/iaa/PCUTS COMB.pep:\*
- 6: /cgn2\_6/prodata/2/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4390	100.0	835	3	US-09-231-839-2
2	4390	100.0	835	4	US-09-458-457-2
3	4390	100.0	835	4	US-09-947-199A-2
4	4079	92.9	835	4	US-09-458-457-8
5	4079	92.9	835	4	US-09-947-199A-8
6	465	10.6	1584	4	US-09-457-040B-27
7	435	9.9	1839	2	US-09-172-977-4
8	435	9.9	1839	4	US-09-404-108-4
9	433.5	9.9	994	4	US-10-164-595-38
10	427.5	9.7	1088	3	US-09-082-059-2
11	419	9.5	843	2	US-09-172-977-3
12	419	9.5	843	4	US-09-404-108-3
13	418	9.5	579	4	US-09-529-279-4
14	418	9.5	579	4	US-10-158-895-4
15	418	9.5	590	4	US-09-529-279-15
16	418	9.5	590	4	US-10-158-895-15
17	404.5	9.2	821	1	US-07-928-464-2
18	404.5	9.2	821	1	US-08-003-311B-2
19	404.5	9.2	821	1	US-08-261-432-2
20	404.5	9.2	821	5	PCT-US93-07347-2
21	400.5	9.1	452	3	US-09-035-706-2
22	400.5	9.1	452	3	US-08-955-841-2
23	400.5	9.1	452	4	US-09-390-425-2
24	400.5	9.1	452	4	US-09-566-906-2
25	397	9.0	1745	2	US-09-031-485-33
26	397	9.0	1745	2	US-08-847-429A-33
27	397	9.0	1745	3	US-09-065-474-33

ALIGNMENTS

RESULT 1

US-09-291-839-2

; Sequence 2, Application US/09291839A

; Patent No. 6261818

; GENERAL INFORMATION:

; APPLICANT: Raju, Jeyaseelan

; TITLE OF INVENTION: NOVEL CARP PROTEIN AND NUCLEIC ACID MOLECULES AND USES THEREFOR

; FILE REFERENCE: WNI-068

; CURRENT APPLICATION NUMBER: US/09/291,839A

; CURRENT FILING DATE: 1999-04-14

; NUMBER OF SEQ ID NOS: 6

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 2

; LENGTH: 835

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-291-839-2

28	397	9.0	1745	4	US-09-557-034-33	Sequence 33, Appl
29	389	8.9	263	3	US-09-035-706-5	Sequence 5, Appl
30	389	8.9	263	3	US-08-955-841-5	Sequence 5, Appl
31	389	8.9	263	4	US-09-390-425-5	Sequence 5, Appl
32	389	8.9	263	4	US-09-566-906-5	Sequence 5, Appl
33	382.5	8.7	352	3	US-09-065-474-139	Sequence 139, App
34	382.5	8.7	352	4	US-09-557-034-139	Sequence 139, App
35	377	8.6	455	3	US-09-221-235-5	Sequence 5, Appl
36	377	8.6	455	3	US-09-221-928-5	Sequence 5, Appl
37	377	8.6	455	3	US-09-221-527-5	Sequence 5, Appl
38	377	8.6	455	3	US-09-221-236-5	Sequence 5, Appl
39	377	8.6	455	3	US-09-221-416-5	Sequence 5, Appl
40	377	8.6	455	3	US-09-221-245-5	Sequence 5, Appl
41	377	8.6	455	3	US-09-163-115-5	Sequence 5, Appl
42	377	8.6	455	3	US-09-221-528-5	Sequence 5, Appl
43	377	8.6	455	3	US-09-593-553-5	Sequence 5, Appl
44	377	8.6	455	3	US-09-221-237-5	Sequence 5, Appl
45	377	8.6	455	4	US-09-399-588-2	Sequence 2, Appl

Query Match 100.0%; Score 4390; DB 3; Length 835;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 835; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MGNYKSRPTQCTDEWKKVSESYVITIERLEDDLOIKEKELTELNRNIFGDEAFSKVNL	60
Db	1	MGNYKSRPTQCTDEWKKVSESYVITIERLEDDLOIKEKELTELNRNIFGDEAFSKVNL	60
Qy	61	NYKTENGLSLHLCCICGGKKSHIRTLMLKGLRPSRLTRNGFTALHLAVYKDNALITSL	120
Db	61	NYKTENGLSLHLCCICGGKKSHIRTLMLKGLRPSRLTRNGFTALHLAVYKDNALITSL	120
Qy	121	LHSGADIQGVGGTALTALHIAIAGHLEAADVLLQHGAVNIQDAVFTPLHIAAYGHE	180
Db	121	LHSGADIQGVGGTALTALHIAIAGHLEAADVLLQHGAVNIQDAVFTPLHIAAYGHE	180
Qy	181	QVTRLLKFGADVNSGVEGDRPLHLASAKGFLNTAKLMBEGSKADVNAQDNEDHVLPH	240
Db	181	QVTRLLKFGADVNSGVEGDRPLHLASAKGFLNTAKLMBEGSKADVNAQDNEDHVLPH	240
Qy	241	FCSRFGHHIVKYLQSDLEVPVNVITGDTPLHLACVNGKFEVAKELIISGTESLTK	300
Db	241	FCSRFGHHIVKYLQSDLEVPVNVITGDTPLHLACVNGKFEVAKELIISGTESLTK	300
Qy	301	ENIFSTAFHSACTYKSIDLVKFLDQNVININHQGRDGHGTGLHSACYHGHIRLVQFL	360
Db	301	ENIFSTAFHSACTYKSIDLVKFLDQNVININHQGRDGHGTGLHSACYHGHIRLVQFL	360
Qy	361	DGADNMLVACDPSSRSSEKDEQTCIMWAYEKGHDAIVTLLKHKYKRPDELPCEYSGP	420
Db	361	DGADNMLVACDPSSRSSEKDEQTCIMWAYEKGHDAIVTLLKHKYKRPDELPCEYSGP	420

QY	421	GDGSYVSPSLGKIKMTKEKADILLRAGLPSHFLQLSIEFHEIIGSGSGKVKYG	480	Db	241	FCRSFGHHDIVKYLQSDLEVPQHVNIYGDTPHLACYNGKFEVAKELIIGSGTESLTK	300
Db	421	GDGSYVSPSLGKIKMTKEKADILLRAGLPSHFLQLSIEFHEIIGSGSGKVKYG	480	QY	301	ENIFSETAFSACTYKGSIDLKFLDQNVININHGDRGHTGLHSACVHGHIRLVQFLL	360
QY	481	RCRNKIYAIKRYRANTYCSKSDVDMFCREVSIICQLNHPCVIOFVGACLNDRPSQFAIVTQ	540	Db	301	ENIFSETAFSACTYKGSIDLKFLDQNVININHGDRGHTGLHSACVHGHIRLVQFLL	360
Db	481	RCRNKIYAIKRYRANTYCSKSDVDMFCREVSIICQLNHPCVIOFVGACLNDRPSQFAIVTQ	540	QY	361	DNGADNMLVACDPSRSRSGEKDEQTCIMWAYEKGHDAIVTLLKHYPKRPDELPCNYSQPG	420
QY	541	YISGSLFSLHHEQKRIIDLOSCLIIVADVAKGMEYLHNLTQPIIHRDLNSHNILLYEDG	600	Db	361	DNGADNMLVACDPSRSRSGEKDEQTCIMWAYEKGHDAIVTLLKHYPKRPDELPCNYSQPG	420
Db	541	YISGSLFSLHHEQKRIIDLOSCLIIVADVAKGMEYLHNLTQPIIHRDLNSHNILLYEDG	600	QY	421	GDGSYVSPSLGKIKMTKEKADILLRAGLPSHFLQLSIEFHEIIGSGSGKVKYG	480
QY	601	HAVVADFGESRFLQSLDEDNTKQGNLRWMAPEVFTQCTRYTIKADVFSYALCLWEILT	660	Db	421	GDGSYVSPSLGKIKMTKEKADILLRAGLPSHFLQLSIEFHEIIGSGSGKVKYG	480
Db	601	HAVVADFGESRFLQSLDEDNTKQGNLRWMAPEVFTQCTRYTIKADVFSYALCLWEILT	660	QY	481	RCRNKIYAIKRYRANTYCSKSDVDMFCREVSIICQLNHPCVIOFVGACLNDRPSQFAIVTQ	540
QY	661	GEIPFAHLKPAADAAADWAYHRIPIGYSIPKPISSLLIRGNWACPEGRPESEVVMKLE	720	Db	481	RCRNKIYAIKRYRANTYCSKSDVDMFCREVSIICQLNHPCVIOFVGACLNDRPSQFAIVTQ	540
Db	661	GEIPFAHLKPAADAAADWAYHRIPIGYSIPKPISSLLIRGNWACPEGRPESEVVMKLE	720	QY	541	YISGSLFSLHHEQKRIIDLOSCLIIVADVAKGMEYLHNLTQPIIHRDLNSHNILLYEDG	600
QY	721	ECLNIELMSPASSNSGSLSPSSSDCLVNRGGPGRSHVAALRSRFELEYALNARSYAA	780	Db	541	YISGSLFSLHHEQKRIIDLOSCLIIVADVAKGMEYLHNLTQPIIHRDLNSHNILLYEDG	600
Db	721	ECLNIELMSPASSNSGSLSPSSSDCLVNRGGPGRSHVAALRSRFELEYALNARSYAA	780	QY	601	HAVVADFGESRFLQSLDEDNTKQGNLRWMAPEVFTQCTRYTIKADVFSYALCLWEILT	660
QY	781	LSQSAGQYSSQGLSLEEMKRSLOQTPIDKYGVSDPMSMHFSCRNSSSFDSS	835	Db	601	HAVVADFGESRFLQSLDEDNTKQGNLRWMAPEVFTQCTRYTIKADVFSYALCLWEILT	660
Db	781	LSQSAGQYSSQGLSLEEMKRSLOQTPIDKYGVSDPMSMHFSCRNSSSFDSS	835	QY	661	GEIPFAHLKPAADAAADWAYHRIPIGYSIPKPISSLLIRGNWACPEGRPESEVVMKLE	720
RESULT 2							
US-09-458-457-2							
; Sequence 2, Application US/09458457							
; Patent No. 6500654							
; GENERAL INFORMATION:							
; APPLICANT: Raju, Jeyaseelan							
; TITLE OF INVENTION: NOVEL CARC PROTEIN AND NUCLEIC ACID MOLECULES AND USES							
; FILE REFERENCE: mni-068cp							
; CURRENT APPLICATION NUMBER: US/09/458.457							
; EARLIER FILING DATE: 1999-12-10							
; EARLIER FILING DATE: 1998-12-11							
; EARLIER FILING DATE: 1999-04-14							
; NUMBER OF SEQ ID NOS: 9							
; SOFTWARE: PatentIn Ver. 2.0							
; SEQ ID NO 2							
; LENGTH: 835							
; TYPE: PRT							
; ORGANISM: Homo sapiens							
US-09-458-457-2							
Query Match 100.0%; Score 4390; DB 4; Length 835;							
Best Local Similarity 100.0%; Pred. No. 0;							
Matches 835; Conservative 0; Mismatches 0; Indels 0; Gaps 0;							
QY	1	MGNYSRPTQCTCTDEWKVKVSESVITIERLEDLQIKEKELTELNRNIFGSDFAFSKVL	60	Db	1	MGNYSRPTQCTCTDEWKVKVSESVITIERLEDLQIKEKELTELNRNIFGSDFAFSKVL	60
Db	1	MGNYSRPTQCTCTDEWKVKVSESVITIERLEDLQIKEKELTELNRNIFGSDFAFSKVL	60	QY	61	NYRTENGLSLLHLCICCGKSKSHIRTLMLKGRPSRLTRNGFTALHLAVYKDAELITSL	120
QY	61	NYRTENGLSLLHLCICCGKSKSHIRTLMLKGRPSRLTRNGFTALHLAVYKDAELITSL	120	Db	61	NYRTENGLSLLHLCICCGKSKSHIRTLMLKGRPSRLTRNGFTALHLAVYKDAELITSL	120
Db	61	NYRTENGLSLLHLCICCGKSKSHIRTLMLKGRPSRLTRNGFTALHLAVYKDAELITSL	120	QY	121	LHSGADIQVGGTALHIATIAHLEAADVLLQHGANNVIOAVFTPLHIAAYGHE	180
QY	121	LHSGADIQVGGTALHIATIAHLEAADVLLQHGANNVIOAVFTPLHIAAYGHE	180	Db	121	LHSGADIQVGGTALHIATIAHLEAADVLLQHGANNVIOAVFTPLHIAAYGHE	180
Db	121	LHSGADIQVGGTALHIATIAHLEAADVLLQHGANNVIOAVFTPLHIAAYGHE	180	QY	181	QVTRLLKFGADVNSGVDRLPLHAKGFLNIAKLLMEEGSKADVNQDNEDHVPLH	240
QY	181	QVTRLLKFGADVNSGVDRLPLHAKGFLNIAKLLMEEGSKADVNQDNEDHVPLH	240	Db	181	QVTRLLKFGADVNSGVDRLPLHAKGFLNIAKLLMEEGSKADVNQDNEDHVPLH	240
Db	181	QVTRLLKFGADVNSGVDRLPLHAKGFLNIAKLLMEEGSKADVNQDNEDHVPLH	240	QY	241	FCRSFGHHDIVKYLQSDLEVPQHVNIYGDTPHLACYNGKFEVAKELIIGSGTESLTK	300
QY	241	FCRSFGHHDIVKYLQSDLEVPQHVNIYGDTPHLACYNGKFEVAKELIIGSGTESLTK	300	Db	241	FCRSFGHHDIVKYLQSDLEVPQHVNIYGDTPHLACYNGKFEVAKELIIGSGTESLTK	300

QY 61 NYRTENGLSLHLCCICGGKSHIRIIMLKGRLPSRLTRNGFTALHLAVYKONAEIITSL 120  
DB 61 NYRTENGLSLHLCCICGGKSHIRIIMLKGRLPSRLTRNGFTALHLAVYKONAEIITSL 120  
QY 121 LHSADIQOQVGYGGLTALHIAATAGHLEAADVLQHGANNVQDAVFFPLHIAAYGHE 180  
DB 121 LHSADIQOQVGYGGLTALHIAATAGHLEAADVLQHGANNVQDAVFFPLHIAAYGHE 180  
QY 181 QVTRLLKFGADVNVSEVGDRLPLHLASAKGFLNIAKLLMEEGSKADVNAQDNEDHVPILH 240  
DB 181 QVTRLLKFGADVNVSEVGDRLPLHLASAKGFLNIAKLLMEEGSKADVNAQDNEDHVPILH 240  
QY 241 FCSRFHGHDIIVKLLQSDLEVOHVNNIYDTPHLHACVNGKFEVAKETIOISGTESITK 300  
DB 241 FCSRFHGHDIIVKLLQSDLEVOHVNNIYDTPHLHACVNGKFEVAKETIOISGTESITK 300  
QY 301 ENIFSETAFHSACTYKSIDLVKFLDQNVNININHQGRDGTGLHSACVHGHIRLVQFLL 360  
DB 301 ENIFSETAFHSACTYKSIDLVKFLDQNVNININHQGRDGTGLHSACVHGHIRLVQFLL 360  
QY 361 DNGADNMLVACDPSRSGEKDEQTCIMWAYEKGHDAIVTLLKHYPQDELPCEYISQPG 420  
DB 361 DNGADNMLVACDPSRSGEKDEQTCIMWAYEKGHDAIVTLLKHYPQDELPCEYISQPG 420  
QY 421 GDGSYVSPPLGKIKSMTKEKADILLRAGLPSHFLQLSIEFHEIIGSGSGFKVYKG 480  
DB 421 GDGSYVSPPLGKIKSMTKEKADILLRAGLPSHFLQLSIEFHEIIGSGSGFKVYKG 480  
QY 481 RCNKIIVAKRYRANTYCSKSDVDMFCREVSILOLNHPCVIOFVGACINDPSPQFAIVTQ 540  
DB 481 RCNKIIVAKRYRANTYCSKSDVDMFCREVSILOLNHPCVIOFVGACINDPSPQFAIVTQ 540  
QY 541 YISGGSIFSLHLEQKRIIDLOSKLIIAVDVAKGMEYLHNLTOPIIHRDLNSHNILLYEDG 600  
DB 541 YISGGSIFSLHLEQKRIIDLOSKLIIAVDVAKGMEYLHNLTOPIIHRDLNSHNILLYEDG 600  
QY 601 HAVVADFGESRFLQSLDEDNMTKQGNLRWMAPEVFTQCTRYTIKADVFSYALCLWEILT 660  
DB 601 HAVVADFGESRFLQSLDEDNMTKQGNLRWMAPEVFTQCTRYTIKADVFSYALCLWEILT 660  
QY 661 GEIPFAHLKFAAAADAMAYHHIRPPIGYSIPKIPISLLIRGNWACPEGRPFSEVVMKLE 720  
DB 661 GEIPFAHLKFAAAADAMAYHHIRPPIGYSIPKIPISLLIRGNWACPEGRPFSEVVMKLE 720  
QY 721 ECLCNIELMSPASSNSGSLSPSSSDCLVNRGGPGRSHVAALRSRFELEYALNARSYAA 780  
DB 721 ECLCNIELMSPASSNSGSLSPSSSDCLVNRGGPGRSHVAALRSRFELEYALNARSYAA 780  
QY 781 LSQSAQYSSQGLSLEEMKRSLOYTPIDKYGYVSDPMSMHFSCRNSSSPEDSS 835  
DB 781 LSQSAQYSSQGLSLEEMKRSLOYTPIDKYGYVSDPMSMHFSCRNSSSPEDSS 835

## RESULT 4

US-09-458-457-8

; Sequence 8, Application US/09458457  
; Patent No. 6500654  
; GENERAL INFORMATION:  
; APPLICANT: Raju, Jeyaseelan  
; TITLE OF INVENTION: NOVEL CARC PROTEIN AND NUCLEIC ACID MOLECULES AND USES  
; FILE REFERENCES: mni-068cp  
; CURRENT APPLICATION NUMBER: US/09/458,457  
; CURRENT FILING DATE: 1999-12-10  
; EARLIER FILING DATE: 1998-12-11  
; EARLIER FILING DATE: 1998-12-11  
; EARLIER FILING DATE: 1999-04-14  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: Patent.in Ver. 2.0  
; SEQ ID NO 8  
; LENGTH: 835  
; TYPE: PRT

## RESULT 5

US-09-947-199A-8

; Sequence 8, Application US/09947199A  
; Patent No. 6660490  
; GENERAL INFORMATION:  
; APPLICANT: Raju, Jeyaseelan  
; TITLE OF INVENTION: NOVEL CARC PROTEIN AND NUCLEIC ACID MOLECULES AND USES  
; FILE REFERENCES: THEREFOR

; ORGANISM: Rattus norvegicus

US-09-458-457-8

Query Match 92.9%; Score 4079; DB 4; Length 835;

Best Local Similarity 91.4%; Pred. No. 0;

Matches 763; Conservative 38; Mismatches 34; Indels 0; Gaps 0;

QY 1 MGNYKSRPTOTCTDEWKXKVSSEYVITIERLEDDQIQEKEKELTELRNIFGSDAFSKVNL 60  
DB 1 MGNYKSRPTOTCTDEWKXKVSSEYVITIERLEDDQIQEKEKELTELRNIFGSDAFSKVNL 60  
QY 61 NYRTENGLSLHLCCICGGKSHIRIIMLKGRLPSRLTRNGFTALHLAVYKONAEIITSL 120  
DB 61 NYRTENGLSLHLCCICGGKSHIRIIMLKGRLPSRLTRNGFTALHLAVYKONAEIITSL 120  
QY 121 LHSADIQOQVGYGGLTALHIAATAGHLEAADVLQHGANNVQDAVFFPLHIAAYGHE 180  
DB 121 LHSADIQOQVGYGGLTALHIAATAGHLEAADVLQHGANNVQDAVFFPLHIAAYGHE 180  
QY 181 QVTRLLKFGADVNVSEVGDRLPLHLASAKGFLNIAKLLMEEGSKADVNAQDNEDHVPILH 240  
DB 181 QVTRLLKFGADVNVSEVGDRLPLHLASAKGFLNIAKLLMEEGSKADVNAQDNEDHVPILH 240  
QY 241 FCSRFHGHDIIVKLLQSDLEVOHVNNIYDTPHLHACVNGKFEVAKETIOISGTESITK 300  
DB 241 FCSRFHGHDIIVKLLQSDLEVOHVNNIYDTPHLHACVNGKFEVAKETIOISGTESITK 300  
QY 301 ENIFSETAFHSACTYKSIDLVKFLDQNVNININHQGRDGTGLHSACVHGHIRLVQFLL 360  
DB 301 ENIFSETAFHSACTYKSIDLVKFLDQNVNININHQGRDGTGLHSACVHGHIRLVQFLL 360  
QY 361 DNGADNMLVACDPSRSGEKDEQTCIMWAYEKGHDAIVTLLKHYPQDELPCEYISQPG 420  
DB 361 DNGADNMLVACDPSRSGEKDEQTCIMWAYEKGHDAIVTLLKHYPQDELPCEYISQPG 420  
QY 421 GDGSYVSPPLGKIKSMTKEKADILLRAGLPSHFLQLSIEFHEIIGSGSGFKVYKG 480  
DB 421 GDGSYVSPPLGKIKSMTKEKADILLRAGLPSHFLQLSIEFHEIIGSGSGFKVYKG 480  
QY 481 RCNKIIVAKRYRANTYCSKSDVDMFCREVSILOLNHPCVIOFVGACINDPSPQFAIVTQ 540  
DB 481 RCNKIIVAKRYRANTYCSKSDVDMFCREVSILOLNHPCVIOFVGACINDPSPQFAIVTQ 540  
QY 541 YISGGSIFSLHLEQKRIIDLOSKLIIAVDVAKGMEYLHNLTOPIIHRDLNSHNILLYEDG 600  
DB 541 YISGGSIFSLHLEQKRIIDLOSKLIIAVDVAKGMEYLHNLTOPIIHRDLNSHNILLYEDG 600  
QY 601 HAVVADFGESRFLQSLDEDNMTKQGNLRWMAPEVFTQCTRYTIKADVFSYALCLWEILT 660  
DB 601 HAVVADFGESRFLQSLDEDNMTKQGNLRWMAPEVFTQCTRYTIKADVFSYALCLWEILT 660  
QY 661 GEIPFAHLKFAAAADAMAYHHIRPPIGYSIPKIPISLLIRGNWACPEGRPFSEVVMKLE 720  
DB 661 GEIPFAHLKFAAAADAMAYHHIRPPIGYSIPKIPISLLIRGNWACPEGRPFSEVVMKLE 720  
QY 721 ECLCNIELMSPASSNSGSLSPSSSDCLVNRGGPGRSHVAALRSRFELEYALNARSYAA 780  
DB 721 ECLCNIELMSPASSNSGSLSPSSSDCLVNRGGPGRSHVAALRSRFELEYALNARSYAA 780  
QY 781 LSQSAQYSSQGLSLEEMKRSLOYTPIDKYGYVSDPMSMHFSCRNSSSPEDSS 835  
DB 781 LSQSAQYSSQGLSLEEMKRSLOYTPIDKYGYVSDPMSMHFSCRNSSSPEDSS 835







QY 432 LGKIKSMTEK-ADIL 446  
Db 661 LLSHLAAEEDKYNVADIL 679

RESULT 12  
US-09-404-108-3  
; Sequence 3, Application US/09404108  
; Patent No. 6590077  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Y. Tom  
; APPLICANT: Guegler, Karl J.  
; APPLICANT: Corley, Neil C.  
; APPLICANT: Yue, Henry  
; TITLE OF INVENTION: HUMAN ANKYRIN FAMILY PROTEIN  
; CURRENT APPLICATION NUMBER: US/09/404,108  
; CURRENT FILING DATE: 1999-09-23  
; EARLIER APPLICATION NUMBER: 09/172,977  
; EARLIER FILING DATE: 1998-10-14  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: PERL Program  
; SEQ ID NO 3  
; LENGTH: 843  
; TYPE: PRT  
; ORGANISM: Rattus norvegicus  
; FEATURE: -  
; OTHER INFORMATION: g1841966  
US-09-404-108-3

Query Match 9.5%; Score 419; DB 4; Length 843;  
Best Local Similarity 29.2%; Pred. No. 1.2e-31;  
Matches 128; Conservative 62; Mismatches 181; Indels 68; Gaps 10;  
QY 60 LNYRTGSLHLLCCIGCGKSHITMLKGLRPSRLTRNGFTALHLAVYKDNALITTS 119  
Db 257 IDAKTRDGLTPLH-CAARSHDQWELLERGAPELLARTKGLSLPLHAAQGDHVECVKH 315  
QY 120 LLSHGADIQGVGGTALHIAIAGHLEAADVLLQHGANNVQDAVFTPLHIAAYGH 179  
Db 316 LLOHKAPVDVTLTYLTALHVAACHGYRVTKLLDKRANPNARALNGFTPLHIAACKNR 375  
QY 180 EQVTRLLKFGADVNSGVEGDEPLHLASAKGFLNIKLLMEGSKAD- 227  
Db 376 IKWELLVKGAVIOAITESGLTPVPAFPMGHNLIVLLQNGASPDVTNIRGETALHM 435  
QY 228 -----VNAODNEDHVLPHFCRFGHHIDIVKYLQSLDLEVPVHVI 268  
Db 436 AARAGEVEVVRCLLRNGALVDARAREEQPLHIASRLGKTEIVQLLQH--MAHPDAATT 493  
QY 269 YGDTPLHLACYNGKFEVAKETIQISTESTLTKENIFSETAFHSACTYKSLDKVFLDQ 328  
Db 494 NGTTPHLHISAREGQVDVSVLLLEAGAAHSLATKGF--TPLHVAAYKG-SLDVAKLLQR 550  
QY 329 NVININHQQRDGTGLHSACYGHIRLVQFLLDNGAMNLVACDPSRSSGKDEQTCILMW 388  
Db 551 RAA-ADSAGKNGLTPLHVAHYDNQKVALLLLEKASPHATA-----KNGYTPHLH 600  
QY 389 AYEKGHDAIVTLKH-----KRPDELPCNEYQPG-----GGSYVSVSP 431  
Db 601 AAKKNQWQIASTLLNLTGAETNTVTKQGVTPHLASQEGHDMVTLVLEKGANIHMSTKSG 660  
QY 432 LGKIKSMTEK-ADIL 446  
Db 661 LLSHLAAEEDKYNVADIL 679

RESULT 13  
US-09-529-279-4  
; Sequence 4, Application US/09529279  
; Patent No. 6451617  
; GENERAL INFORMATION:

; APPLICANT: ONO, KOICHIRO  
; APPLICANT: OHTOMO, TOSHIHIKO  
; APPLICANT: TSUCHIYA, MASAYUKI  
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES  
; FILE REFERENCE: 053466/0278  
; CURRENT APPLICATION NUMBER: US/09/529,279  
; CURRENT FILING DATE: 2000-04-11  
; PRIOR APPLICATION NUMBER: PCT/JP98/04796  
; PRIOR FILING DATE: 1998-10-22  
; PRIOR APPLICATION NUMBER: JP 9/290188  
; PRIOR FILING DATE: 1997-10-22  
; NUMBER OF SEQ ID NOS: 48  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 579  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-529-279-4  
Query Match 9.5%; Score 418; DB 4; Length 579;  
Best Local Similarity 31.6%; Pred. No. 7.5e-32;  
Matches 115; Conservative 67; Mismatches 144; Indels 38; Gaps 13;

QY 458 LQLSEIEFHEIIGSGFGKVKGRCKNKIVAKRYRANTYCKSDVDMFCEVSVILCOLN 517  
Db 31 IDYKEIEVEEVGGRGAFGVVCKAKWRAKDVAIKQIE-----SESEKAPIVELROLSEVN 85  
QY 518 HPCVIOFVGACLNDRPSQFAIVTYIISGSLFSLH--EOKRILDLQSKLIIIAVDVAKGME 575  
Db 86 HPNIVKLYGACLN---PVCLVMEYAEGLSLYVHLGAERPLPYTAAHAMSNCLOCSQGVA 142  
QY 576 YLHNLTOP--IHRDLNLSNILLYDGHAV-VADGESRFLQSLDEDNMTQPGNLRWMA 632  
Db 143 YLSHM-OPKALHHRDLKPNLLVAGTVLKICDFGTACDIQT-----HMTNNKGSAAWMA 197  
QY 633 PEVFTOCTRYTKADVFSVALCWLBEILTGEIPFAHLKPAADAAADWAYHH-IRPPIGYGIP 691  
Db 198 PEVPE-EGSYSEKCDVFSWGIILNEVITRKEPFDIGGPAPRIMWAVENGTRPPLIKNLP 256  
QY 692 KPISLLILRGWNAPEGRPEFSEVYVVKLEECI-----CNIELMSPA-----SSNS 737  
Db 257 KPISLMTKCSKDPQRSORPMEEIVKIMTHLMRYFFGADPELPQYPCQYSDGQSNATST 316  
QY 738 GSLSPSSSDCLVNRGPGSRSHVALRSERFELEYALNARSYAALSQSAGQYS---SQGLS 794  
Db 317 GSPMDIATSN-TSNKSDTNNMQVPATNDTIKRLSKLKNQAKQOSESGRUSLGASHGSS 375  
QY 795 LEM 798  
Db 376 VESL 379

RESULT 14  
US-10-158-895-4  
; Sequence 4, Application US/10158895  
; Patent No. 6551840  
; GENERAL INFORMATION:  
; APPLICANT: ONO, KOICHIRO  
; APPLICANT: OHTOMO, TOSHIHIKO  
; APPLICANT: TSUCHIYA, MASAYUKI  
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES  
; FILE REFERENCE: 053466/0278  
; CURRENT APPLICATION NUMBER: US/10/158,895  
; CURRENT FILING DATE: 2002-06-03  
; PRIOR APPLICATION NUMBER: US/09/529,279  
; PRIOR FILING DATE: 2000-04-11  
; PRIOR APPLICATION NUMBER: PCT/JP98/04796  
; PRIOR FILING DATE: 1998-10-22  
; PRIOR APPLICATION NUMBER: JP 9/290188  
; PRIOR FILING DATE: 1997-10-22  
; NUMBER OF SEQ ID NOS: 48  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 4

QY 576 YLHNLTOP--IHRDLNSHNILLYEDGHAV-VADFGESRFLQSLDEDNMTKOPGNLRWMA 632  
Db 143 YLHSM-OPKALIHRLDKPPNLLVAGGTVLKICDFGTACDIQT-----HMTNKGSAANWA 197  
QY 633 PEVFTQCTRYTIKADVFSYALCLWEILTGEIIPFAHLKPAADAAADVAYHH-IRPPIGYSIP 691  
Db 198 PEVF-EGSNYSKCDVFSNGIILWEVITRRKPFDEIGGPAFRIMWAVHNGTRPPLIKNLP 256  
QY 692 KPISSLLIRGNWACPEGRPEFSEVVMKLECL-----CNIELMSPA-----SSNSS 737  
Db 257 KPIESLMTRCWSDQSPQSPMSBEIVKIMTHLMRYFPFGADEPLOYPCQYSDGQSNATST 316  
QY 738 GSLSPSSSDCLVNRGGPGRSHVAALRSRFELEYALNARSYAALSQSAGQYS---SQGLS 794  
Db 317 GSFWDIASTN-TSNKSDTNMEQVPATNDTIKRLSKLLKNQAKQOQSESGRLSLGASHGSS 375  
QY 795 LEEM 798  
Db 376 VESL 379

Search completed: September 2, 2004, 14:25:45  
Job time : 48 secs

QY 458 LQSLSEIEFHEIIGSGFGKVKGRCKRNIKVAIKRYRANTYCSKSDVDMFCREVSILCOLN 517  
Db 31 IDYKEIEVEEVVGRGAFGVCKAKWRAKDVAIKQIE-----SESERKAFIVELRQLSRVN 85  
QY 518 HPCVIOFVAGCLNDPSQFAIVTQYISGSLFSLH--EQRKILDQSKLIIADVAKGME 575  
Db 86 HPNIVKLYGACLN---PVCLVMEYAEAGGSLYNVLHGAEPDPYTTAAHAMSWCLQCQSGVA 142  
QY 576 YLHNLTOP--IHRDLNSHNILLYEDGHAV-VADFGESRFLQSLDEDNMTKOPGNLRWMA 632  
Db 143 YLHSM-OPKALIHRLDKPPNLLVAGGTVLKICDFGTACDIQT-----HMTNKGSAANWA 197  
QY 633 PEVFTQCTRYTIKADVFSYALCLWEILTGEIIPFAHLKPAADAAADVAYHH-IRPPIGYSIP 691  
Db 198 PEVF-EGSNYSKCDVFSNGIILWEVITRRKPFDEIGGPAFRIMWAVHNGTRPPLIKNLP 256  
QY 692 KPISSLLIRGNWACPEGRPEFSEVVMKLECL-----CNIELMSPA-----SSNSS 737  
Db 257 KPIESLMTRCWSDQSPQSPMSBEIVKIMTHLMRYFPFGADEPLOYPCQYSDGQSNATST 316  
QY 738 GSLSPSSSDCLVNRGGPGRSHVAALRSRFELEYALNARSYAALSQSAGQYS---SQGLS 794  
Db 317 GSFWDIASTN-TSNKSDTNMEQVPATNDTIKRLSKLLKNQAKQOQSESGRLSLGASHGSS 375  
QY 795 LEEM 798  
Db 376 VESL 379

RESULT 15  
US-09-529-279-15  
; Sequence 15, Application US/09529279  
; Patent No. 6451617  
; GENERAL INFORMATION:  
; APPLICANT: ONO, KOICHIRO  
; APPLICANT: OHTOMO, TOSHIHIKO  
; APPLICANT: TSUCHIYA, MASAYUKI  
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES  
; FILE REFERENCE: 053466/0278  
; CURRENT APPLICATION NUMBER: US/09/529,279  
; CURRENT FILING DATE: 2000-04-11  
; PRIOR APPLICATION NUMBER: PCT/JF98/04796  
; PRIOR FILING DATE: 1998-10-22  
; PRIOR APPLICATION NUMBER: JP 9/290188  
; PRIOR FILING DATE: 1997-10-22  
; NUMBER OF SEQ ID NOS: 48  
; SOFTWARE: PatentIn ver. 2.1  
; SEQ ID NO 15  
; LENGTH: 590  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-529-279-15

Query Match 9.5%; Score 418; DB 4; Length 590;  
Best Local Similarity 31.6%; Pred. No. 7.8e-32;  
Matches 115; Conservative 67; Mismatches 144; Indels 38; Gaps 13;  
QY 458 LQSLSEIEFHEIIGSGFGKVKGRCKRNIKVAIKRYRANTYCSKSDVDMFCREVSILCOLN 517  
Db 31 IDYKEIEVEEVVGRGAFGVCKAKWRAKDVAIKQIE-----SESERKAFIVELRQLSRVN 85  
QY 518 HPCVIOFVAGCLNDPSQFAIVTQYISGSLFSLH--EQRKILDQSKLIIADVAKGME 575  
Db 86 HPNIVKLYGACLN---PVCLVMEYAEAGGSLYNVLHGAEPDPYTTAAHAMSWCLQCQSGVA 142

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 2, 2004, 14:24:19 ; Search time 142 Seconds

(without alignments)

1852.672 Million cell updates/sec

Title: US-10-626-173-2

Perfect score: 4390

Sequence: 1 MGNYSRPTQTCTDEWKVY.....PMSSMHFHSRNSSEFSDS 835

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1298764 seqs, 315065143 residues

Total number of hits satisfying chosen parameters: 1298764

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4390	100.0	835	9	US-09-947-199-2
2	4390	100.0	835	16	US-10-626-173-2
3	4079	92.9	835	9	US-09-947-199-8
4	4079	92.9	835	16	US-10-626-173-8
5	1893	43.1	359	16	US-10-664-421-138
6	486.5	11.1	359	16	US-10-437-963-195691
7	483.5	11.0	376	12	US-10-424-599-271998
8	482	11.0	603	16	US-10-437-963-156923
9	480.5	10.9	478	12	US-10-425-114-40404
10	478.5	10.9	585	12	US-10-425-114-67797
11	473	10.8	323	12	US-10-425-114-69530
12	473	10.8	462	12	US-10-425-114-67647
13	472.5	10.8	338	12	US-10-425-114-64160
14	468	10.7	551	12	US-10-424-599-196246
15	466.5	10.6	422	12	US-10-425-114-45517

16	464.5	10.6	357	12	US-10-424-599-252964
17	459.5	10.5	346	12	US-10-425-114-57043
18	458.5	10.4	545	16	US-10-437-963-195284
19	457	10.4	427	12	US-10-424-599-193555
20	457	10.4	424	12	US-10-425-114-36844
21	452	10.3	376	16	US-10-437-963-175612
22	442	10.1	422	16	US-10-437-963-132604
23	440.5	10.0	1053	12	US-10-221-278-343
24	440.5	10.0	1053	15	US-10-291-172-343
25	439.5	10.0	483	12	US-10-424-599-284054
26	439.5	10.0	469	12	US-10-425-114-49887
27	439.5	10.0	469	12	US-10-425-114-56056
28	438.5	10.0	3993	16	US-10-437-963-168323
29	438	10.0	429	12	US-10-425-114-57924
30	438	10.0	429	12	US-10-425-114-67484
31	437.5	10.0	483	12	US-10-425-114-40870
32	436.5	9.9	919	15	US-10-108-260A-4122
33	435.5	9.9	453	12	US-10-424-599-284055
34	434.5	9.9	542	16	US-10-437-963-196087
35	433	9.8	454	12	US-10-425-114-40125
36	432	9.8	425	9	US-09-828-313-29
37	432	9.8	1330	15	US-10-108-260A-3237
38	431.5	9.8	417	12	US-10-425-114-46822
39	431	9.8	459	12	US-10-425-114-72504
40	430	9.8	604	16	US-10-437-963-173220
41	428	9.7	482	16	US-10-437-963-189054
42	427.5	9.7	367	16	US-10-250-613-6
43	426	9.7	536	16	US-10-437-963-149523
44	423.5	9.6	453	12	US-10-424-599-149508
45	423	9.6	426	12	US-10-425-114-60778

#### ALIGNMENTS

#### RESULT 1

US-09-947-199-2  
; Sequence 2, Application US/09947199  
; Patent No. US20020127684A1  
; GENERAL INFORMATION:  
; APPLICANT: Raju, Jeyaseelan  
; TITLE OF INVENTION: NOVEL CARC PROTEIN AND NUCLEIC ACID MOLECULES AND USES  
; TITLE OF INVENTION: THEREFOR  
; FILE REFERENCE: MNI-068CP2  
; CURRENT APPLICATION NUMBER: US/09/947,199  
; CURRENT FILING DATE: 2001-09-05  
; PRIOR APPLICATION NUMBER: 60/111,938  
; PRIOR FILING DATE: 1998-12-11  
; PRIOR APPLICATION NUMBER: 09/291,839  
; PRIOR FILING DATE: 1999-04-14  
; PRIOR APPLICATION NUMBER: 09/458,457  
; PRIOR FILING DATE: 1999-12-10  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 835  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; US-09-947-199-2

Query Match 100.0%; Score 4390; DB 9; Length 835;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 835; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MGNYSRPTQTCTDEWKVYVITIERLEDDLOIKEKELTELNRNIFGDEAFSKVNL 60  
Db 1 MGNYSRPTQTCTDEWKVYVITIERLEDDLOIKEKELTELNRNIFGDEAFSKVNL 60  
Qy 61 NYRTENGLSLHLCICCGGKSHIRTLMLKGLRPSRLTNGFTALHLVYKDNELITSL 120  
Db 61 NYRTENGLSLHLCICCGGKSHIRTLMLKGLRPSRLTNGFTALHLVYKDNELITSL 120  
Qy 121 LHSAGDIQGVGGTALHIATIAHLEAADVLLQGANVNIQDAVFTPLHIAAYGHE 180

121	Db	LHS	GADI	QOV	GGGL	TALH	AT	IAGH	LEAD	VL	LQ	HGAN	VNI	OD	VFT	PL	HI	AA	YGH	180
181	QY	QVTR	LL	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	240
181	QY	QVTR	LL	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	240
241	QY	FC	SR	G	H	H	I	V	K	L	L	L	L	L	L	L	L	L	L	300
241	QY	FC	SR	G	H	H	I	V	K	L	L	L	L	L	L	L	L	L	L	300
301	QY	EN	I	F	S	E	T	A	F	H	S	A	C	T	Y	G	S	I	L	360
301	QY	EN	I	F	S	E	T	A	F	H	S	A	C	T	Y	G	S	I	L	360
361	QY	D	N	G	A	D	M	N	L	V	A	C	D	P	S	R	S	S	E	420
361	QY	D	N	G	A	D	M	N	L	V	A	C	D	P	S	R	S	S	E	420
421	QY	G	D	G	S	V	Y	V	P	S	P	L	G	K	I	K	S	M	T	480
421	QY	G	D	G	S	V	Y	V	P	S	P	L	G	K	I	K	S	M	T	480
481	QY	R	C	R	N	I	V	A	I	K	R	Y	A	N	T	Y	C	S	K	540
481	QY	R	C	R	N	I	V	A	I	K	R	Y	A	N	T	Y	C	S	K	540
541	QY	Y	I	S	G	S	L	F	S	L	L	H	E	Q	R	I	L	D	L	600
541	QY	Y	I	S	G	S	L	F	S	L	L	H	E	Q	R	I	L	D	L	600
601	QY	H	A	V	A	D	P	G	E	S	R	F	L	S	D	E	N	T	K	660
601	QY	H	A	V	A	D	P	G	E	S	R	F	L	S	D	E	N	T	K	660
661	QY	G	E	I	P	F	A	H	L	K	P	A	A	A	A	A	D	M	Y	720
661	QY	G	E	I	P	F	A	H	L	K	P	A	A	A	A	A	D	M	Y	720
721	QY	E	C	L	N	I	E	L	M	S	P	A	S	S	N	S	G	S	L	780
721	QY	E	C	L	N	I	E	L	M	S	P	A	S	S	N	S	G	S	L	780
781	QY	L	S	Q	A	G	Y	S	S	Q	G	L	S	E	E	M	K	R	S	835
781	QY	L	S	Q	A	G	Y	S	S	Q	G	L	S	E	E	M	K	R	S	835

## RESULT 2

```

US-10-626-173-2
; Sequence 2, Application US/10626173
; Publication No. US20040110232A1
; GENERAL INFORMATION:
; APPLICANT: Raju, Jeyaseelan
; TITLE OF INVENTION: NOVEL CARK PROTEIN AND NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR

```

Query Match	100.0%;	Score 4390;	DB 16;	Length 835;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 835;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MGNYSKSP	TQCTDEWKKY	SESVITIERLEDLQIKEKELTELNRNIGSDEAFSKVNL 60
DB	1	MGNYSKSP	TQCTDEWKKY	SESVITIERLEDLQIKEKELTELNRNIGSDEAFSKVNL 60
QY	61	NYRTENGLS	LLHLCCICGGK	KSHIRIMUKGRPSRLTRNGFTALHLAVYKDNABLIITSL 120
DB	61	NYRTENGLS	LLHLCCICGGK	KSHIRIMUKGRPSRLTRNGFTALHLAVYKDNABLIITSL 120
QY	121	LHSGADTQO	VCYGGTALHTAT	TAGHLEADVLLQHCANVNIQDAVFTPLHIAAYGYHE 180
DB	121	LHSGADTQO	VCYGGTALHTAT	TAGHLEADVLLQHCANVNIQDAVFTPLHIAAYGYHE 180
QY	181	QVTRLLLKFG	ADVNVSGEVD	RPLHLASAKGFNIAKLLMEBSGSKADVNAQDNEDHVPJH 240
DB	181	QVTRLLLKFG	ADVNVSGEVD	RPLHLASAKGFNIAKLLMEBSGSKADVNAQDNEDHVPJH 240
QY	241	FCSPFGHDI	IVKYLQSDLE	VQPHVNIYGDTPHLACYNGKFEVAKELIQISGTESITK 300
DB	241	FCSPFGHDI	IVKYLQSDLE	VQPHVNIYGDTPHLACYNGKFEVAKELIQISGTESITK 300
QY	301	ENIETSETAF	SACTYGKSID	LVKFLDQNVININHQGRDGTGLHSACYVGHIRIVQFLL 360
DB	301	ENIETSETAF	SACTYGKSID	LVKFLDQNVININHQGRDGTGLHSACYVGHIRIVQFLL 360
QY	361	DNGADMNL	LVACDPSR	SGEKDEQTCMLWAEYKEGHDAIVTLLKHYPQDELP
DB	361	DNGADMNL	LVACDPSR	SGEKDEQTCMLWAEYKEGHDAIVTLLKHYPQDELP
QY	421	GDGSYVSV	PSPLGKIKMT	KEKADILLRAGLPSHFHLQISEIEFHEIIGSGSFQKVKY 480
DB	421	GDGSYVSV	PSPLGKIKMT	KEKADILLRAGLPSHFHLQISEIEFHEIIGSGSFQKVKY 480
QY	481	RCRNKIVA	IKRYRANTY	CYSKSDVDMFCREVSIICQLNHP
DB	481	RCRNKIVA	IKRYRANTY	CYSKSDVDMFCREVSIICQLNHP
QY	541	YISGGSFSL	HEOKRILDI	QSKLITAVDVAKGWEYLHNLTOPIIHRDLNSHNILLYEDG 600
DB	541	YISGGSFSL	HEOKRILDI	QSKLITAVDVAKGWEYLHNLTOPIIHRDLNSHNILLYEDG 600
QY	601	HAVVADGES	RFLQSLDED	NNTKQGNLRWMAPEVTTQCTRYTIKADVFSYALCWEILIT 660
DB	601	HAVVADGES	RFLQSLDED	NNTKQGNLRWMAPEVTTQCTRYTIKADVFSYALCWEILIT 660
QY	661	GEIPFHLK	PAAAADWAY	HHIRPPIGYSIPKPISSILLIRGNWACPEGP
DB	661	GEIPFHLK	PAAAADWAY	HHIRPPIGYSIPKPISSILLIRGNWACPEGP
QY	721	ECLCNIELM	S	PASSNSGSLSPSSSDCLVNRGPGGRSHVAALRSRFFLEYALNARSYAA 780
DB	721	ECLCNIELM	S	PASSNSGSLSPSSSDCLVNRGPGGRSHVAALRSRFFLEYALNARSYAA 780
QY	781	LSQAGQY	SSQGLSIE	EMKRSLOYPIDKTYGYSDPMSSMHFSCRNSSSEFSSS 835
DB	781	LSQAGQY	SSQGLSIE	EMKRSLOYPIDKTYGYSDPMSSMHFSCRNSSSEFSSS 835

### RESULT 3

RESOLUT 3  
 US-09-947-199-8  
 ; Sequence 8, Application US/09947199  
 ; Patent No. US20020127684A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Raju, Jevaseelan  
 ; TITLE OF INVENTION: NOVEL CARX PROTEIN AND NUCLEIC ACID MOLECULES AND USES  
 ; TITLE OF INVENTION: THEREFOR  
 ; FILE REFERENCE: MNI-068CP2  
 ; CURRENT APPLICATION NUMBER: US/09/947,199  
 ; CURRENT FILING DATE: 2001-09-05

```
; PRIOR APPLICATION NUMBER: 60/111,938
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/291,839
; PRIOR FILING DATE: 1999-04-14
; PRIOR APPLICATION NUMBER: 09/458,457
; PRIOR FILING DATE: 1999-12-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 835
; TYPE: PRT
; ORGANISM: Rattus norvegicus
; US-09-947-199-8

Query Match          92.9%; Score 4079; DB 9; Length 835;
Best Local Similarity 91.4%; Pred. No. 2.4e-309;
Matches 763; Conservative 38; Mismatches 34; Indels 0; Gaps 0;

QY 1 MGNYSRPTCTCDEWKKVSESYVITIERLEDDLOIKEKELTELNRNIFGSDFAFSKYNL 60
DB 1 MGNYSRPTCTCDEWKKVSESYVITIERLEDDLOIKEKELTELNRNIFGSDFAFSKYNL 60
QY 61 NYRTENGLSLHLCCCTCGGKKSHIRLMKGLRPSRLTRNGFTALHLAVYKDNABLIITSL 120
DB 61 NYRTENGLSLHLCCCTCGGKKSHIRLMKGLRPSRLTRNGFTALHLAVYKDNABLIITSL 120
QY 121 LHSAGADIQQVGYGGLTALHIAITAGHLEAADVLLQHGANNVQDAVFTPLHIAAYYGH 180
DB 121 LHSAGADVQQVGYGGLTALHIAITAGHPEAAEVLLQHGANNVQDAVFTPLHIAAYYGH 180
QY 181 QVTRLLKFGADVNVSGEVDRLPLHLASAKGFLNIAKLMEEGSKADVNAQDNEDHVP 240
DB 181 QVTSVLLKFGADVNVSGEVDRLPLHLASAKGFFNIVKLLVEEGSKADVNAQDNEDHVP 240
QY 241 FCSRFGHHDIKVKLLQSDLEVPQHVNYIGDTPHLACVNGKFEVAKELIQTGTSITK 300
DB 241 FCSRFGHHNIVSYLLQSDLEVPQHVNYIGDTPHLACVNGKFEVAKELIQTGTSITK 300
QY 301 ENIFSETAFHSACTYKSIDLVKFLDDQNVININHGDRDGTGLHSACYHGHIRLVQFLL 360
DB 301 ENIFSETAFHSACTYKSIDLVKFLDDQNVININHGDRDGTGLHSACYHGHIRLVQFLL 360
QY 361 DNGADNMLVACDPSRSGKEQOTCLMWAYEKGHDAIVTLLKHYPKQDELPCHNEYSQ 420
DB 361 DNGADNMLVACDPSRSGKEQOTCLMWAYEKGHDAIVTLLKHYPKQDELPCHNEYSQ 420
QY 421 GDGSYVSVPSPLGKI KSMTEKADILLRAGLPSHFHLQLSIEPHEIIGSGSGKVKY 480
DB 421 GDGSYVSVPSPLGKI KSMTEKADILLRAGLPSHFHLQLSIEPHEIIGSGSGKVKY 480
QY 481 RCNKKIVAKRYRANTYCSKSDVDMFCREVSVILCOLNHPCVIQFVACLNPPSQAIV 540
DB 481 RCNKKIVAKRYRANTYCSKSDVDMFCREVSVILCOLNHPCVIQFVACLNPPSQAIV 540
QY 541 YISGGSLSLLHEQKRILOSKLIIAVDVAKGMEYHLNLTQPIIHRDLNSHNILLYED 600
DB 541 YISGGSLSLLHEQKRILOSKLIIAVDVAKGMEYHLNLTQPIIHRDLNSHNILLYED 600
QY 601 HAVVADFGESRFLQSILDEDNMTKQGNLRWMAPEVFTQCTRYTIKADVFYALCLWE 660
DB 601 HAVVADFGESRFLQSILDEDNMTKQGNLRWMAPEVFTQCTRYTIKADVFYALCLWE 660
QY 661 GEIPEAHLKPAADAAADMAVHHIRPPIGYGSIKPIISLLIRGNACPEGRPFSEVW 720
DB 661 GEIPEAHLKPAADAAADMAVHHIRPPIGYGSIKPIISLLIRGNACPEGRPFSEVW 720
QY 721 ECLCNIELMSPASSNSGSLSPSSSDCLLVNRGGPGRSHVAALRRFRFEYALNARS 780
DB 721 ECLCNIELMSPASSNSGSLSPSSSDCLLVNRGGPGRSHVAALRRFRFEYALNARS 780
QY 781 LSGSAGYSSQGLSLEMKRSIQYTPIDKYGVSDPMGSMHPSCRNSSFEDSS 835
DB 781 WSQSVGTHNPGLSLEEMNRSTQYSTVDKYGVSDPMGLTHLSRQDDNSFEDSN 835
```

```
; PRIOR APPLICATION NUMBER: 60/111,938
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/291,839
; PRIOR FILING DATE: 1999-04-14
; PRIOR APPLICATION NUMBER: 09/458,457
; PRIOR FILING DATE: 1999-12-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 835
; TYPE: PRT
; ORGANISM: Rattus norvegicus
; US-09-947-199-8
```

```
RESULT 4
US-10-626-173-8
; Sequence 8, Application US/10626173
; Publication No. US20040110232A1
; GENERAL INFORMATION:
; APPLICANT: Raju, Jeyaseelan
; TITLE OF INVENTION: NOVEL CARC PROTEIN AND NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: MNI-068CP2
; CURRENT APPLICATION NUMBER: US/10/626,173
; CURRENT FILING DATE: 2003-07-24
; PRIOR APPLICATION NUMBER: US/09/947,199A
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: US 09/291,839
; PRIOR FILING DATE: 1999-04-14
; PRIOR APPLICATION NUMBER: US 09/458,457
; PRIOR FILING DATE: 1999-12-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 835
; TYPE: PRT
; ORGANISM: Rattus norvegicus
; US-10-626-173-8
```

```
Query Match          92.9%; Score 4079; DB 16; Length 835;
Best Local Similarity 91.4%; Pred. No. 2.4e-309;
Matches 763; Conservative 38; Mismatches 34; Indels 0; Gaps 0;

QY 1 MGNYSRPTCTCDEWKKVSESYVITIERLEDDLOIKEKELTELNRNIFGSDFAFSKYNL 60
DB 1 MGNYSRPTCTCDEWKKVSESYVITIERLEDDLOIKEKELTELNRNIFGSDFAFSKYNL 60
QY 61 NYRTENGLSLHLCCCTCGGKKSHIRLMKGLRPSRLTRNGFTALHLAVYKDNABLIITSL 120
DB 61 NYRTENGLSLHLCCCTCGGKKSHIRLMKGLRPSRLTRNGFTALHLAVYKDNABLIITSL 120
QY 121 LHSAGADIQQVGYGGLTALHIAITAGHLEAADVLLQHGANNVQDAVFTPLHIAAYYGH 180
DB 121 LHSAGADVQQVGYGGLTALHIAITAGHPEAAEVLLQHGANNVQDAVFTPLHIAAYYGH 180
QY 181 QVTRLLKFGADVNVSGEVDRLPLHLASAKGFLNIAKLMEEGSKADVNAQDNEDHVP 240
DB 181 QVTSVLLKFGADVNVSGEVDRLPLHLASAKGFFNIVKLLVEEGSKADVNAQDNEDHVP 240
QY 241 FCSRFGHHDIKVKLLQSDLEVPQHVNYIGDTPHLACVNGKFEVAKELIQTGTSITK 300
DB 241 FCSRFGHHNIVSYLLQSDLEVPQHVNYIGDTPHLACVNGKFEVAKELIQTGTSITK 300
QY 301 ENIFSETAFHSACTYKSIDLVKFLDDQNVININHGDRDGTGLHSACYHGHIRLVQFLL 360
DB 301 ENIFSETAFHSACTYKSIDLVKFLDDQNVININHGDRDGTGLHSACYHGHIRLVQFLL 360
QY 361 DNGADNMLVACDPSRSGKEQOTCLMWAYEKGHDAIVTLLKHYPKQDELPCHNEYSQ 420
DB 361 DNGADNMLVACDPSRSGKEQOTCLMWAYEKGHDAIVTLLKHYPKQDELPCHNEYSQ 420
QY 421 GDGSYVSVPSPLGKI KSMTEKADILLRAGLPSHFHLQLSIEPHEIIGSGSGKVKY 480
DB 421 GDGSYVSVPSPLGKI KSMTEKADILLRAGLPSHFHLQLSIEPHEIIGSGSGKVKY 480
QY 481 RCNKKIVAKRYRANTYCSKSDVDMFCREVSVILCOLNHPCVIQFVACLNPPSQAIV 540
DB 481 RCNKKIVAKRYRANTYCSKSDVDMFCREVSVILCOLNHPCVIQFVACLNPPSQAIV 540
QY 541 YISGGSLSLLHEQKRILOSKLIIAVDVAKGMEYHLNLTQPIIHRDLNSHNILLYED 600
DB 541 YISGGSLSLLHEQKRILOSKLIIAVDVAKGMEYHLNLTQPIIHRDLNSHNILLYED 600
QY 601 HAVVADFGESRFLQSILDEDNMTKQGNLRWMAPEVFTQCTRYTIKADVFYALCLWE 660
```

Db 601 HAVVADGESRFQSLDEDNMTKOPGNLRWMAEVFTQCTRYTIKADVFSYSLCWELLT 660  
Qy 661 GEIPFAHLKPAADAAADWYHHPPIGYSIPKPSISLLIRGNWACPEGRPESEVVMKLE 720  
Db 661 GEIPFAHLKPAADAAADWYHHPPIGYSIPKPSISLLIRGNWACPEGRPESEVVMKLE 720  
Qy 721 ECLCNIELMSPASSNSGSLSPSSSDCLVNRGPGRSHVAALRSRFELEYALNARSAA 780  
Db 721 ECLCNIELMSPASSNSGSLSPSSSDCLVNRGPGRSHVAALRSRFELEYALNARSAA 780  
Qy 781 LSQAGQYSSOGLSLEEMKRSLOVTPIDKYGYVSDPMSSMHFSCNRSNSSFDDSS 835  
Db 781 WSQSVGTHSNPGLSLEEMRSTQYSTVDKYGYVSDPMSLTHLSRQDDSNFSDSN 835

## RESULT 5

US-10-664-421-138  
; Sequence 138, Application US/10664421  
; Publication No. US200401428641

## GENERAL INFORMATION:

; APPLICANT: BREWER, RYAN  
; APPLICANT: IBRAHIM, PRAHA  
; APPLICANT: KUMAR, ABINAV  
; APPLICANT: MANDIYAN, VALSAN  
; APPLICANT: MILBURN, MICHAEL V.  
; TITLE OF INVENTION: CRYSTAL STRUCTURE OF PIM-1 KINASE  
; FILE REFERENCE: 039363/0703  
; CURRENT APPLICATION NUMBER: US/10/664,421  
; CURRENT FILING DATE: 2003-09-16

; PRIOR APPLICATION NUMBER: 60/412,341

; PRIOR FILING DATE: 2002-09-20

; PRIOR APPLICATION NUMBER: 60/411,398

; PRIOR FILING DATE: 2002-09-16

; NUMBER OF SEQ ID NOS: 169

; SOFTWARE: PatentIn Ver. 3.2

; SEQ ID NO 138

; LENGTH: 359

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-664-421-138

Query Match 43.1%; Score 1893; DB 16; Length 359;

Best Local Similarity 100.0%; Pred. No. 3.4e-139; Mismatches 0; Indels 0; Gaps 0;  
Matches 359; Conservative 0;

Qy 413 CNEYSQGGDGSYVSPPLGKIKSMTEKADILLRAGLPSPHFHQLQSLSEIETIIGSG 472  
Db 1 CNEYSQGGDGSYVSPPLGKIKSMTEKADILLRAGLPSPHFHQLQSLSEIETIIGSG 60  
Qy 473 SFGKYKGRCNKIIVAKRYRANTYCSKSDVDMFCREVSIILCOLNHPCVIQVGCINDP 532  
Db 61 SFGKYKGRCNKIIVAKRYRANTYCSKSDVDMFCREVSIILCOLNHPCVIQVGCINDP 120  
Qy 533 SQAFAIVTOYISGGSLFSLHHEQKRIIDLOSKLIIVADVAKGMEYHLNLTQPIIHRDLNSH 592  
Db 121 SQAFAIVTOYISGGSLFSLHHEQKRIIDLOSKLIIVADVAKGMEYHLNLTQPIIHRDLNSH 180  
Qy 593 NLLYEDGHAVVADGESRFLQSLDEDNMTKQGNLRWMAEVFTQCTRYTIKADVFSYA 652  
Db 181 NLLYEDGHAVVADGESRFLQSLDEDNMTKQGNLRWMAEVFTQCTRYTIKADVFSYA 240  
Qy 653 LCLWEILLTGEPFAHLKPAADAAADWYHHPPIGYSIPKPSISLLIRGNWACPEGRPEF 712  
Db 241 LCLWEILLTGEPFAHLKPAADAAADWYHHPPIGYSIPKPSISLLIRGNWACPEGRPEF 300  
Qy 713 SEVVMKLECLCNIELMSPASSNSGSLSPSSSDCLVNRGPGRSHVAALRSRFELEY 771  
Db 301 SEVVMKLECLCNIELMSPASSNSGSLSPSSSDCLVNRGPGRSHVAALRSRFELEY 359

## RESULT 6

US-10-437-963-195691

; Sequence 195691, Application US/10437963  
; Publication No. US20040123343A1

## GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping

; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53221)B

; CURRENT APPLICATION NUMBER: US/10/437,963

; CURRENT FILING DATE: 2003-05-14

; NUMBER OF SEQ ID NOS: 204966

; SEQ ID NO 195691

; LENGTH: 359

; TYPE: PRT

; ORGANISM: Oryza sativa

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_91615C.1.pap

US-10-437-963-195691

Query Match 11.1%; Score 486.5; DB 16; Length 359;

Best Local Similarity 37.0%; Pred. No. 2.1e-29; Mismatches 105; Indels 37; Gaps 9;  
Matches 117; Conservative 57;

Qy 468 IIGSGFGKYKGRCNKIIVAKRYRANTYCSKSDVDMFCREVSIILCOLNHPCVIQVGA 527  
Db 49 MIASGCGDLYHGTLYGEDVAVKILRSE-HLKNVWNEFTQEVILREVQHTNVVRFGA 107  
Qy 528 CLNDPSQFAIVTOYISGGSLFSLHHEQKRIIDLOSKLIIVADVAKGMEYHLNLTQPIIHR 587  
Db 108 C-TKPPQFCITTEYMSGGSLYDFVHRQHNVLDTLLKFAVDVCRGMCYLRQ--RGIHR 164  
Qy 588 DLNHNILLYEDGHAVVADGESRFLQSLDEDNMTKQGNLRWMAPE-VFTQCT----- 640  
Db 165 DLKSNLLMDKHVVAVDFGVARFQD--QGGNMTAETGYRWMAPEHFFSPCSLGERMN 222  
Qy 641 -----RYTIKADVFSYALCLWEILLTGEPFAHLKPAADAAADWYHHPPIGYS 689  
Db 223 WKKEEIVINHQPYDNKADVFSFAIVLWELITSKIPTYNTMTPLQAAVGVV--QGLRFGDPEN 281  
Qy 690 IPRPISLLIRGNWACPEGRPESEVVMKLECLCNIELMSPASS-----NSSGSLSPSS 744  
Db 282 AHPOLLDMRCWEGIPSNRPPFPFDILAEILLARVQEQEHSNQMKHEHSSSNLK--- 338  
Qy 745 SSDCLVNRGPGRSHV 760  
Db 339 ----LSNAGMLGSNYI 350

## RESULT 7

US-10-424-599-271998

; Sequence 271998, Application US/10424599

; Publication No. US20040031072A1

## GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J  
; APPLICANT: Kovalic, David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53223)B

; CURRENT APPLICATION NUMBER: US/10/424,599

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 285684

; SEQ ID NO 271998

; LENGTH: 376

; TYPE: PRT

; ORGANISM: Glycine max

```
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(376)
; OTHER INFORMATION: unsure at all xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_87633C.1.p.p
US-10-424-599-271998

Query Match      11.0%; Score 483.5; DB 12; Length 376;
Best Local Similarity 38.8%; Pred. No. 3.9e-29;
Matches 112; Conservative 54; Mismatches 108; Indels 15; Gaps 7;

Qy 456 FHLQSBIEBHEIIGSGFGKVKGRCKNIKAIKRYRANTYCSKSDVDM---FCREYSI 512
Db 93 WEIDTQKXENKVGSGFDLVRGYCSQDVAIKVLP-----ERISTDLREFAEYVI 148
Qy 513 LCOLNHPCVIQFGACLNPDSPFAIVTOYISGSLFSLHHEQKRLDLOSKLIIIAVDVAK 572
Db 149 MRKIRHNKVVOQFIGACTRPN-LCIVTFEWSRGLYDFLHKQGVFKLPSSLKVAIDVSK 207
Qy 573 GMEYLNLTQPIIHRDLNSHNLIIYEDGHAVVADFGESRFLQSLDEDNMTKQGNLRMWA 632
Db 208 GMYLHQ--NNIHRDLKTANLLMDENEVVKVADFGVARV--QTSQGVNTASTGYRMA 263
Qy 633 PEVFTQCTRTIKADVFSALCLWEILTGEIPPAHLKPAADAAADMAVHHRIPPIGYSIPK 692
Db 264 PEVI-EHKPYDQKADVFSFGIALWELLTGELPYSCLTPLQAAVGVVQKGLRPTIPKNTHP 322
Qy 693 PISGLLRGNACPEGRPESEVVMKLEBCLNIELMSPASSNSGSL 741
Db 323 RLSELLQRCWQDPTORPNFSEIIEILOQIA--KEYNDHKDKSSHGFLS 369

RESULT 8
US-10-437-963-156923
; Sequence 156923, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Brad
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 156923
; LENGTH: 603
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_56544C.1.p.p
US-10-437-963-156923

Query Match      11.0%; Score 482; DB 16; Length 603;
Best Local Similarity 29.4%; Pred. No. 9.8e-29;
Matches 130; Conservative 83; Mismatches 189; Indels 40; Gaps 10;

Qy 296 ESLTKENIFSETAFPHSACTYCKSIDLVKFLLDONVININHQGRD-----GHTGLHSA 347
Db 154 EEVASRLNLNQSSHPFPFAGSSTNLALALEAS----KSQGDHSDSDNVNRPMEI 209
Qy 348 CYHGHIR-----LVQFLLDNGADMLNVLACDPSSSGEKDEQCLMWAYEKHDAIVILL 401
Db 210 TFSIDRPKLISELTSLLGLGLNIQEAHAFSTNDGYSLDFVVGWGHDEETDLIESVR 269
Qy 402 KHYKRPQDELPCNEYSQPGDGGSVSVPSPLGKIKMTKEKADILLRAGLPSHFHLQLS 461

; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(376)
; OTHER INFORMATION: unsure at all xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_87633C.1.p.p
US-10-424-599-271998

Query Match      11.0%; Score 483.5; DB 12; Length 376;
Best Local Similarity 38.8%; Pred. No. 3.9e-29;
Matches 112; Conservative 54; Mismatches 108; Indels 15; Gaps 7;

Qy 456 FHLQSBIEBHEIIGSGFGKVKGRCKNIKAIKRYRANTYCSKSDVDM---FCREYSI 512
Db 93 WEIDTQKXENKVGSGFDLVRGYCSQDVAIKVLP-----ERISTDLREFAEYVI 148
Qy 513 LCOLNHPCVIQFGACLNPDSPFAIVTOYISGSLFSLHHEQKRLDLOSKLIIIAVDVAK 572
Db 149 MRKIRHNKVVOQFIGACTRPN-LCIVTFEWSRGLYDFLHKQGVFKLPSSLKVAIDVSK 207
Qy 573 GMEYLNLTQPIIHRDLNSHNLIIYEDGHAVVADFGESRFLQSLDEDNMTKQGNLRMWA 632
Db 208 GMYLHQ--NNIHRDLKTANLLMDENEVVKVADFGVARV--QTSQGVNTASTGYRMA 263
Qy 633 PEVFTQCTRTIKADVFSALCLWEILTGEIPPAHLKPAADAAADMAVHHRIPPIGYSIPK 692
Db 264 PEVI-EHKPYDQKADVFSFGIALWELLTGELPYSCLTPLQAAVGVVQKGLRPTIPKNTHP 322
Qy 693 PISGLLRGNACPEGRPESEVVMKLEBCLNIELMSPASSNSGSL 741
Db 323 RLSELLQRCWQDPTORPNFSEIIEILOQIA--KEYNDHKDKSSHGFLS 369

RESULT 9
US-10-425-114-40404
; Sequence 40404, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 40404
; LENGTH: 478
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB143-022-C7_FLI.p.p
US-10-425-114-40404

Query Match      10.9%; Score 480.5; DB 12; Length 478;
Best Local Similarity 27.4%; Pred. No. 9.3e-29;
Matches 170; Conservative 95; Mismatches 179; Indels 177; Gaps 29;

Qy 145 GHLEAADVLLQHGANNIQAIVFFFTPLHIAAYGHEQVTRLLKFGADVNVSGEVDRLP 204
Db 7 GYAEAD-----LEPRD-----RLAAHFG-----RLPHSYQLDINVD-KASDVIV 45
Qy 205 H---LASAKG-----FLNIAKLMEEGSKADVNAQDNEDHVLPHFCRSRFGH-HDI 250
Db 46 HONVLAEKDPDRPAPYVRFRLRIEDMSAYSDASEEGDDGDDLSVRQDTPYTHIHEI 105
Qy 251 VKYLLQSDLEVOQPHVNVNYGDTPLHLACYNKGFVAKETIIQISGTES-----LTKENIFS 305
Db 106 VFSTI-----DKP-----KLSQSLALLSDIGLNIRAHVFS 137
Qy 306 ETAFHSACTYCKSIDLVKFLLDONVININHQGRDGTGLHSA CYEHGIRLVQFLLDNGAD 365
Db 138 T---HD---GYSLDV--FVWDGWPV---QDAD---GLHKA-----163
Qy 366 MNLVACDPSSSGEKDEQCLMWAYEKHDAIVILLKHYKRPQDELPCNEYSQPGDGGSY 425
Db 164 ---LEASILRNEGS-----WSGS-----SHSSAAERTLP---FOVKGGEW---197
```



```

; SEQ ID NO 67647
; LENGTH: 462
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3354-074-A5_FLI.pep
US-10-425-114-67647

Query Match 10.8%; Score 473; DB 12; Length 462;
Best Local Similarity 38.7%; Pred. No. 3.4e-28;
Matches 108; Conservative 57; Mismatches 96; Indels 18; Gaps 8;

QY 468 ITGSGFGKVKYGRCKNIVAKIRYRANTYCSKSDVDMFCREVSILCOLNHPCVIQFVGA 527
DB 193 MVASCSGDLHGTYLGEDVAVKIRAB-HLNKNVWNEFTQEVVILREVQHKVWRFIGA 251
QY 528 CLNDSQFAIVTQYISGSGLSFLSHQKRIIDLQSKJIIIVDVAKGMEYILHNLTQPIIHR 587
DB 252 C-TKPPQFCITTEYMSGSLYDFVHKQHNVLNRLTLKFAVDVCRGMCYLHE--RGIHR 308
QY 588 DLNSINILLYEDGHAVADFESRFLQSLDDN--MTKQPGNLRWMAPEVFTQCTRYTI 644
DB 309 DUKTANLMLKMDKHHVVKVADFVGARF-----QDQGWMTAETGTYRWMAPEVINH-QPYDN 362
QY 645 KADVFYSYALCWEILTGEIPFAHLKPAADMAAYHHRPPIGYSIPKPISSILLIRGWA 704
DB 363 KADVFSAIVWELITSKIPIYESMPTLQAAVGV-RQLRPLGPKKTHPKVLDLMQRCWEA 421
QY 705 CPEGEPSEVVMKLEECLECHIE-----LMSPASSNSGS 739
DB 422 DPSARPAFPDILAELEDLLAQVQGTGKTAHAPSNNSNT 460

RESULT 13
US-10-425-114-64160
; Sequence 64160, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingtong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack B.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 64160
; LENGTH: 338
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3592-081-E1_FLI.pep
US-10-425-114-64160

Query Match 10.8%; Score 472.5; DB 12; Length 338;
Best Local Similarity 37.0%; Pred. No. 2.4e-28;
Matches 108; Conservative 58; Mismatches 113; Indels 13; Gaps 6;

QY 432 LGKIKNTYKADILLRAGLPSHFLQLSEIEFHEIIGSGSGFKGVYKGRCKNIVAKIR 491
DB 26 LKMQIAENSAADRVEIPTDGASEWIEDVKKLFKNKVASGSDGLYRGTCSDQVAIKY 85
QY 492 YRANTYCSKSDVDM--FCREVSILCOLNHPCVIQFVGACLNDPSQFAIVTQYISGSGLSF 548
DB 86 LKP-----ERINADMQRFAEQEYIWRKVRKNVQFQIACCTKPN-ICVITFMSGSGSVY 140
QY 549 SILHEQKRIIDLQSKJIIIVDVAKGMEYILHNLTQPIIHRDLNSHNLILYEDGHAVADF 608
DB 141 DYLRHKHGVKFLPALVGATDVSKGMSYLHO--NNIHRDLKNTANLMDNGTGVKVDG 198

; SEQ ID NO 67647
; LENGTH: 462
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3354-074-A5_FLI.pep
US-10-425-114-67647

Query Match 10.7%; Score 468; DB 12; Length 551;
Best Local Similarity 33.9%; Pred. No. 1.1e-27;
Matches 122; Conservative 64; Mismatches 148; Indels 26; Gaps 8;

QY 381 DEQTCLMWAYEGHDAIVTLLKHVKRPQDBLPCNYSQPGDGSGYVSPGLGKIKSMTK 440
DB 209 EETEKALKALRG-----VLKIERQVKSP-----QSVSSVDKP---DQAKMK 249
QY 441 EKADILLRAGLPSHFLQLSEIEFHEIIGSGSGFKGVYKGRCKNIVAKIRYRANTYCSK 500
DB 250 SELDYLTPTDGTVDWEIDPKHLKYGTQIASGSGYSGELFKGVYQSEVAIKVLKADHVNSE 309
QY 501 SDVDMFCREVSILCOLNHPCVIQFVGACLNDPSQFAIVTQYISGSGLSFLSHQKRIIDL 560
DB 310 LQRE-FAQEVYIMRKVRKNVQFQIACCTKSP-RLCIVTBFMSGSGSYDYLRHKQKGFKE 367
QY 561 QSKLIIIVDVAKGMEYILHNLTQPIIHRDLNSHNLILYEDGHAVADFESRFTQSLDEN 620
DB 368 PLLKVAIDVSKGMYLHQ--HNIHRDLKAAULLMDENCTVAVDFGVARV--KAQSGV 423
QY 621 MTKQPGNLRWMAPEVFTQCTRYTIKADVFYSYALCWEILTGEIPFAHLKPAADMAAYH 680
DB 424 MTAETGTYRWMAPEVI-EHKPYDHKADVFSGFVIGVWELLTGKLPFYELTFLQAAIGVQK 482
QY 681 HIRPPIGYSIPKPISSILLIRGWAPEGRPEFSEVVMKLEECLECNLTLMSPASSNSGSL 740
DB 483 GLRPTIPKNTHPKFVELLERSWQODPLRPFDFSEIITLQQLAKEVGDGERHKDKSGGL 542

RESULT 15
US-10-425-114-45517
; Sequence 45517, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingtong
; APPLICANT: Zhou, Yihua

```

```
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 45517
; LENGTH: 422
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700243456_FLI.pep
US-10-425-114-45517

Query Match          10.6%; Score 466.5; DB 12; Length 422;
Best Local Similarity 32.9%; Pred. No. 9.7e-28;
Matches 131; Conservative 62; Mismatches 142; Indels 63; Gaps 15;

QY 354 RIQVFLDNGADMILVAC-----DPSRSSGGEKDEQTCIMWAYEKGHDAIVTLKHYKRPQ 408
Db 9 RRARELL--SISTNPIACHKVCYTDASHCFNQRSIQL-----HP---RKLVRSGDM 55

QY 409 DELPCNEYSQGGGSGYSVSP-----LQKIKMTKEKADILL-----LRAGLPS 454
Db 56 KRLQCFKSSGGGNGG-----GPGKRLERLSLGEYKKAIVSKYLVAPPGARIRSGGEE 110

QY 455 HFHLQLSEIEFHEIIGSGFGKVKYGRCKNKIVAIAKRYRANTYCSKSDVD-----M 505
Db 111 LWSADLSKLEIRGFASGRHSRVYSGRYTGREVAIK-----WVSGPEEDALAAELERQ 164

QY 506 FCREVSIQLQNLHPCVIQFVGACLNQPSQFAIVTQYISGSLFSLLEOK-RILDLSKL 564
Db 165 FASEVALLLRLHHPNIIISFVAAC-KKPPVFCIITEFMAGGSLRKYLHQEPHSVPLNLVL 223

QY 565 IIAVDVAKGMEVHLNLTQPIIHRDLNSHNILLYEDGHAVVADFGESRFLQSLDEDNMTKQ 624
Db 224 KLALDIARGMYYLH--SQGILHRLKSENILLGEDMSVKVDFG----ISCSLQCGSGK 277

QY 625 --PGNLRWMAPEVFTQCTRYTIKADVFSYALCLWEILTGPIFAHLKPAAADMAVHHI 682
Db 278 GFTGYRWAPFEMIKE-EHHTRKVDVYSFGIVMWEILTALVPFSDMTPEQAAVAVALKVA 336

-QY 683 RPPIGYSIPKPISSILLINGWACPEGRPEFSEVVMKLE 720
Db 337 RPPLPASCFAISHLIWQCWATNPDKRPQFDDIVAIL 374
```

Search completed: September 2, 2004, 14:36:20  
Job time : 146 secs